R E P O R T R E S U M E S

ED 002 824 24

PERCEPTION OF SYMBOLS IN SKILL LEARNING BY MENTALLY RETARDED, GIFTED, AND NORMAL CHILDREN.

BY- HARRIS, THEODORE L. AND OTHERS
WISCONSIN UNIV., MADISON, SCH. OF EDUC.

REPORT NUMBER CRP-263 PUB DATE JUL 61
CONTRACT OEC-SAE-7135
EDRS PRICE MF-$0.36 HC-$8.96 224P.

DESCRIPTORS- *HANDWRITING SKILLS, *COMPARATIVE ANALYSIS, *ELEMENTARY SCHOOL STUDENTS, *MENTAL RETARDATION, WRITING SKILLS, *GIFTED, HANDWRITING INSTRUCTION, MADISON, WISCONSIN

THE PROCESS OF PERCEIVING LANGUAGE SYMBOLS AS CUES FOR MEANING, MODELS FOR SKILL PERFORMANCE, AND DEFINITIONS OF VALUE PERCEPTION WAS INVESTIGATED. THE OBJECTIVES WERE (1) TO COMPARE THE PERFORMANCE ON SELECTED PERCEPTUAL TASKS OF SAMPLING POPULATIONS OF MENTALLY RETARDED, AVERAGE, AND SUPERIOR CHILDREN, AND (2) TO ANALYZE THESE FINDINGS IN RELATION TO THEIR IMPLICATIONS FOR THE TEACHING-TRAINING PROCESS. THIRTY PUPILS IN GRADES 4, 5, AND 6 AT 3 IQ LEVELS--60-69, 96-109, AND 130-139--WERE TESTED AND STUDIED INTENSIVELY FOR PERCEPTION OF FUNDAMENTAL SCHOOL LEARNING TASKS OF WRITING, READING, SPELLING, AND ARITHMETIC. THE PRESENCE AND DIFFERENTIATION OF NORMATIVE AND ASPIRATIONAL PERCEPTIONS HAVE BEEN DEMONSTRATED AS BEING PRESENT IN THE WRITING BEHAVIOR OF ALL INTELLIGENCE GROUPS, AND THE LACK OF STABILITY IN THE WRITING PERFORMANCE OF MIDDLE-GRADE STUDENTS SUGGESTED THE IMPORTANCE OF THIS KIND OF DEVELOPMENT IN THE SKILL TRAINING OF ELEMENTARY SCHOOL CHILDREN. RESEARCH AND REDIRECTION OF SKILL TRAINING ARE NEEDED, AND A MORE ADEQUATE HANDWRITING EDUCATIONAL PROGRAM IS REQUIRED. THIS WAS A CONTINUATION OF PROJECT NO. 151. (GC)

EO002824

3414A

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

PERCEPTION OF SYMBOLS IN SKILL LEARNING BY MENTALLY RETARDED, GIFTED, AND NORMAL CHILDREN

Theodore L. Harris
Virgil E. Herrick
G. Lawrence Rarick

The Committee for Research in Basic Skills
Virgil E. Herrick, Chairman

School of Education
University of Wisconsin

Project SAE 6436 extended to
Project SAE 7135

This research reported herein was performed
pursuant to a contract with
United States Office of Education
Department of Health, Welfare, and Education

July, 1961

FOREWORD

This study addresses itself to the general problem of trying to gain a greater understanding of how children learn the basic skills important in the educational programs of all schools. More specifically, it is concerned with the examination of the perceptual models a child uses to control and improve his handwriting behavior. Handwriting is used as the focus of this investigation because of behavioral availability, it's universality in the school experience of every child, and because of its usefulness in serving as a medium to examine problems which seem to undergird the whole domain of school skill learning. Handwriting, therefore, is regarded as a prototype for research which can easily be extended to such other skill areas as reading, spelling, arithmetic, and composition.

Because of this comprehensive interest in what the handwriting act may manifest, effort was made in this investigation to consider a measure of quality to serve as a dependent variable to check the significance of such other factors and correlates as variation of the handwriting task, writing speed, pressure control, specific and general muscular tension, and the affective state of the individual. The relationship of these data to the different perceptual tasks presented to each subject and to the three conditions of intelligence forms the basis for most of the conclusions of this study.

Every effort was made in this study to maximize the generalizability of findings through careful sampling procedures, control of the testing and experimental conditions in both the laboratory and field investigations, preciseness of the measurement of the experimental variables, and the full reporting of

research procedures. At every point in the investigation the major investigators were involved in the full range of research tasks.

The number of school systems, individual teachers and administrators and children who have participated in this research are too numerous for us to extend our appreciation to them for their cooperation and contribution to our study on an individual basis. It is obvious certainly that their help was essential to the completion of this effort and that no data could have been collected without their participation. To all these individuals, we offer our heartfelt thanks.

The following individuals have served as research assistants on this project: Rodney Paulsen, Mrs. Elizabeth Fennema, Mrs. Lynne Schellberg Hirst, Miss Ella Bullis, Mrs. Gay Bryan, Mrs. Leatrice Anderson Hinz, Mrs. Adrienne Erlebacher, Dan Andersen, Wayne Otto, and Miss Nora Okada.

The Committee acknowledges the many contributions the above group has made to the study in ideas, suggestions for improvement in methodology, and in the interpretation of data. Not the least is our appreciation of their devotion to the values of careful data collection and accuracy in computation. Their willingness to keep the hard work of any research project in perspective and to share the discouragement as well as the victories with humor and good will is much appreciated.

Appreciation is extended to the U. S. Office of Education whose sponsorship and financial support has made possible this project. The assistance of the staff of the U. S. Office on the problems of budget and reporting is fully recognized.

The unfailing assistance and support of the Department of Education, Lindley J. Stiles, Dean of the School of Education, and the Research Committee of the Graduate School through commitments of time, space, and additional grants gave the day to day support which is so necessary to the final realization of a research program of this kind.

Because any project as comprehensive as this draws on related resources, credit should be given to the long time support of The Parker Pen Company of

Janesville, Wisconsin, to the Committee's study of handwriting behavior. Many of the background studies and instrumentation were possible only through their generosity and encouragement.

Mrs. Shirley Otto has served as executive secretary through the major portion of this study and deserves great credit for the able assistance in keeping budget records, supervising of hourly help, filing of data, and in the development and typing of reports and publications. Her role in the final publication of this project report is fully recognized.

The Committee gives final recognition to the many children who have contributed their samples of handwriting and who have tried so hard to live up to all of the demands of the experimental tasks and expectancies. It is only hoped that the knowledge gained from this and subsequent research will make their skill learning more efficient and meaningful.

THE COMMITTEE FOR RESEARCH IN BASIC SKILLS

Theodore L. Harris

Theodore L. Harris
Professor of Education

Virgil E. Herrick

Virgil E. Herrick, Chairman
Professor of Education

G. Lawrence Rarick

G. Lawrence Rarick
Professor of Men and Women's
Physical Education

TABLE OF CONTENTS

	Page
FOREWORD	ii
LIST OF TABLES	vi
LIST OF FIGURES	xi
Chapter	
I. Perception of Symbols in Skill Learning by Mentally Retarded, Gifted, and Normal Children	1
II. The Evaluation of Legibility in Handwriting	13
III. Children's Perception of the Handwriting Task	51
IV. Psychological and Motor Correlates of Handwriting Legibility	98
V. The Relationship Between Handwriting Performance and Independent Measures of Motor Control	142
VI. Conclusions and Discussion	165
APPENDICES	184

LIST OF TABLES

Table	Page
1. Characteristics of Scales for Evaluation of Legibility in Handwriting.	17
2. Distribution of Individuals and Writing Samples: Scaling Project I	27
3. Grade Levels of Handwriting Samples: Scaling Project IV	32
4. Number of Ratings Used in Evaluating Items.	38
5. Indication of Weight of Each of the Four Slant Categories	42
6. Grade Levels of Handwriting Samples: Handwriting Project V.	45
7. Acceptable and Unacceptable Handwriting Samples: Handwriting Project V	45
8. Background Data for Four Studies of the Handwriting of Bright, Average, and Low I.Q. Children.	53
9. Nature and Order of Tasks Used in the Perception of the Handwriting Task.	65
10. Arrangement of Scaled Writing Samples in Order from Easiest to Hardest to Read: Subject 4.	67
11. D^2 Table for Scoring Rearrangement Items: Score (Number of Things Ranked) x $(Rho)^2$.	68
12. Analysis of Variance Among Bright, Average, and Dull Groups in Ordering Handwriting Samples Correctly (Milwaukee).	71
13. "t" Test Results for the Significance of Difference in Means Between Bright, Average, and Dull Groups in Ordering Handwriting Samples Correctly (Milwaukee)	71
14. Number and Per Cent of Complete Agreement in Ordering by Bright, Average, and Dull Groups (Milwaukee).	72
15. Number and Per Cent of Complete Agreement in Ordering by Bright and Average Groups (Madison).	72
16. Number and Per Cent of Near Agreement in Ordering by Bright, Average, and Dull Groups (Milwaukee).	73

Page

17. Number and Per Cent of Near Agreement in Ordering by Bright and Average Children (Madison) 73

18. Agreement on Successive Orderings of "Like Mine" Samples (Milwaukee) 76

19. Agreement on Successive Orderings of "Like Mine" Samples (Madison) 77

20. Agreement on Successive Orderings of "Aspired To" Samples (Milwaukee) 78

21. Agreement on Successive Ordering of "Aspired To: Samples (Madison) 78

22. Relative Position of "Like Mine" and "Aspired To" Samples in the First Ordering of Scale Samples (Milwaukee) 79

23. Relative Position of "Like Mine" and "Aspired To" Samples in the Second Ordering of Scale Samples (Milwaukee) 79

24. Relative Position of "Like Mine" and "Aspired To" Samples in the First Ordering of Scale Samples (Madison) 81

25. Relative Position of "Like Mine" and "Aspired To" Samples in the Second Ordering of Scale Samples (Madison) 81

26. Relative Position in First and Second Ordering of Best Sample (Milwaukee) 81

27. Relative Position in First and Second Ordering of Poorest Sample (Milwaukee) 81

28. Relative Position of Best and Poorest Samples in First Ordering (Milwaukee) 82

29. Relative Position of Best and Poorest Samples in Correct Ordering (Milwaukee) 82

30. Relative Position of Best and Poorest Samples in First Ordering (Madison) 83

31. Relative Position of Best and Poorest Samples in Correct Ordering (Madison) 83

32. Analysis of Variance: Relationship of Legibility to I.Q., Sex, and Conditions (Milwaukee) 84

33. Analysis of Variance: Relationship of Legibility to I.Q., Sex, and Conditions (Madison) 84

34. Correspondence of the Ordering of Bright, Average, and Dull Children of Their Handwriting Samples and Adult Scaling (Milwaukee) 86

Page

35. Correspondence of the Ordering of Bright and Average Children of Their Handwriting Samples for Legibility with Adult Scaling (Madison). 86

36. The Average Difference in Quality Between Two Usual Writing Samples as a Consequence of I.Q. and Sex (Milwaukee). . . . 88

37. The Average Difference in Quality Between Two Usual Writing Samples as a Consequence of I.Q. and Sex (Madison). 88

38. Range of Quality in the Three Conditions of Writing as a Consequence of I.Q. and Sex (Milwaukee). 89

39. Range of Quality in the Three Conditions of Writing as a Consequence of I.Q. and Sex (Madison). 89

40. Range of Legibility of the Writing Samples Best, Normal, and Fast by Intelligence and Sex (Milwaukee). 90

41. Relationship Between Legibility, Perception, and Intelligence (Milwaukee). 91

42. Relationships Between Legibility, Perception, and Intelligence (Madison). 92

43. Means and Standard Deviations of Writing Times for Bright Children and Children of Average I.Q. According to Condition 116

44. Summary of Analysis of Variance for Writing Rates in Bright and Average I.Q. Children 117

45. Intercorrelations Among Writing Rates by Task Conditions for Bright and Average Boys and Girls. 118

46. Means and Standard Deviations of Force Variation Ratios, An Index of Fine Motor Control, for Bright and Average I.Q. Children According to the Condition 119

47. Summary of Analysis of Variance for Fine Motor Control (FVR) in Bright and Average Children. 120

48. Inter-Relationships Among Force Variation Ratios by Task Conditions for Bright and Average Boys and Girls. 121

49. Frequency Tabulation of Ranks for Each of the Writing Tasks According to I.Q. Groups. 122

50. Within Subject Agreement Among the Four Psychophysiological Measures for the Fine Conditions of Writing as Determined by the Coefficient of Concordance:W 123

51. Summary of Analysis of Variance for Legibility in Bright and Average Children. 124

52. Means and Standard Deviations of Legibility Scores of Bright Children and Children of Average Intelligence for Four Conditions . 124

Page

53. Intercorrelations Among Legibility Ratings by Task Conditions for Bright and Average Boys and Girls 126

54. Number and Per Cent of Bright and Average Subjects Producing Samples with Legibility and Force Variation Greater than that of the First Normal Sample. 127

55. Number and Per Cent of Bright and Average Subjects in which an Increase in Legibility is Accompanied by a Decrease in Force Variation 128

56. Legibility, Speed, and Fine Motor Control (FVR) Scores for Selected Good and Poor Writers Among Bright and Average Children. . 129

57. Summary of Friedman Two-Way Analysis of Variance of Rank Order for Each of the Six Factors 132

58. Frequency Tabulation of Cases by Ranks for Each of the Writing Conditions Grouped According to Four Psychophysical Measures and Legibility for Children of Average and High Intelligence. 134

59. Results of the Analysis of Variance on the Circle Motor Control Test for 90 Milwaukee Subjects. 147

60. Means and Standard Deviations on Circle Motor Control Test for 90 Milwaukee Subjects 148

61. Results of the Analysis of Variance on the Circle Motor Control Test for 30 Madison Subjects. 149

62. Means and Standard Deviations on Circle Motor Control Test for 30 Madison Subjects 150

63. Inter-Trial Consistency on the Circle Motor Control Test for 90 Milwaukee Subjects. 151

64. Inter-Trial Consistency on the Circle Motor Control Test for 20 Bright and Average Madison Subjects 152

65. Inter-Trial Consistency for Time Scores on the Circle Motor Control Test for 20 Bright and Average Madison Subjects . . . 152

67. Inter-Trial Consistency in Legibility for 20 Bright and Average Madison Subjects. 154

68. Inter-Trial Consistency in Writing Speed for 90 Milwaukee Subjects. 154

69. Inter-Trial Consistency in Legibility for 90 Milwaukee Subjects. 154

70. Correlations Between Average Scores on Trials 1 - 2 and 3 - 4 on the Circle Motor Control Test and Legibility for 90 Milwaukee Subjects 155

Page

71. Correlations Between Average Scores on Trials 1 - 2 and 3 - 4 on the Circle Motor Control Test and Legibility for 20 Bright and Average Madison Subjects. 156

72. Correlations Between Average Speed Scores on Trials 1 - 2 3 - 4 on the Circle Motor Control Test and Legibility for 20 Bright and Average Madison Subjects. 156

73. Correlations Between Average Scores on Trials 1 - 2 and 3 - 4 on the Circle Motor Control Test and the Force Variation Ratio for 20 Bright and Average Madison Subjects. 157

74. Comparative Means of Hits and Performance Times on the Circle Motor Control Test for 20 Bright and Average Madison and 90 Milwaukee Subjects 158

LIST OF FIGURES

Figure	Page
1. Scale Lines.	35
2. Master Recording Card.	44
3. Diagram of Testing Booth and Placement of Apparatus and Personnel.	109
4. Placement of Dual Metallic Electrodes for Contact for the GSR. . .	110
5. Top of Writing Table with Adjustable Metal Platen.	111
6. Circle Motor Control Test.	145

CHAPTER I

PERCEPTION OF SYMBOLS IN SKILL LEARNING BY MENTALLY RETARDED, GIFTED, AND NORMAL CHILDREN

One of the important tasks of the elementary school in any culture is to help children learn as efficiently and effectively as possible the basic skills which are equally important to the individual learner and to the society of which he is a part. Research directed at improving our knowledge of how these skills are learned more effectively, therefore, has greater potential consequence for affecting more individuals and practices at significant points in educational programs than has work in almost any other problem area in the school domain.

This study, reported as a part of a much more comprehensive program of research in skill learning, is an investigation of how children of differentiated intelligence levels in the middle grades of the elementary school perceive their handwriting. The behavior of these children in perceiving a common skill learning task under experimentally controlled conditions is also examined in relation to qualitative differences in legibility, motor control, point pressure variation, action potentials in relevant muscle systems, and evidences of affective tension as indicated by GSR's. This study with its initial exploratory phase covered the four year period from July, 1957, to July, 1961.

Limits and Rationale of the Study

Examination of the perception of symbols in skill learning by mentally retarded, gifted, and normal children obviously does not take on much meaning until a number of definitions and limitations are made in relation to the important

components of this problem. Delineations will be made of the following components: nature of perception phenomena, nature and limits of the research task, educational period of development considered, nature of physiological and psychological correlates, the sex variable, differentiation as a test of significance, and the sampling theory.

Nature of Perception

It is possible to identify four kinds of perception in verbal skill learning: 1) the perception and differentiation of the visual and auditory configurations of letters, words, and arithmetical symbols; 2) the attachment of meaning to letters, words, and more complex language units; 3) the development of normative perception behavior or set in the development of form and meaning of letters, words, sentences, and other complex verbal skill processes; and 4) the goal setting expectancy (level of aspiration) as determined by a desired quality of performance. The key terms in this statement are differentiation of form, attachment of meaning, normative set, and aspiration model of quality.

This study is limited to the examination of the last two kinds of perception in skill learning--the development of a normative model or set which controls and limits the variability in the skill act and the development of a desired model of quality of skill behavior. This is not to minimize the first two kinds of perception in the skill act. It is merely argued that these two latter phases of perception represent a higher maturity level in the learning of skills, are more critical in their adequate maintenance, and are relatively unexplored research-wise than are the other phases of perception.

Using Allport's[1] terminology, the normative and aspirational perceptions by the learner in the skill act are more concerned with the problems of "assembling and building up" of perceptual structures, their flexibility in distortion

---

[1]Floyd H. Allport, Theories of Perception and the Concept of Structure (New York: John Wiley and Sons, Inc., 1955), pp. 598-605.

and modification, the establishment and persistence of constant relationships, and their energic level maintenance and weighting characteristics. These topics form important referent points in all perceptual and learning theories and thus form a research area in verbal skill learning which has great significance for skill development programs in the elementary school.

The work of Harlow,[2] Spence,[3] Bruner,[4] Postman,[5] and Helson[6] establish for all learning the importance of learning and perceptual set (a generalized ability to learn any discrimination problem), the role that symbolic value has in organizing perception, and the stability of the perceptual set (veridicality) over time. More recently, the 1959 symposium of the American Psychological Association[7] on "A Decade of 'New Look' Perception" summarized the current position on issues in this area. Three trends were noteworthy: 1) less interest in the motivated modification of perception, 2) growing recognition of the essential stability of perceptual structures, and 3) recognition that perception has both cognitive and affective phenomenon.

The recent discussions of Mowrer[8] and Osgood[9] establish the central importance of perceptual phenomena in all learning theory, particularly as it deals

---

[2]H. F. Harlow, "The Formation of Learning Sets," Psychological Review, 56 (1949), 51-65.

[3]K. W. Spence, "The Nature of the Response in Discrimination Learning," Psychological Review, 59 (1952), 89-93.

[4]J. S. Bruner and L. Postman, "Symbolic Value as an Organizing Factor in Perception," Journal of Social Psychology, 27 (1948), 203-208.

[5]L. Postman and R. S. Crutchfield, "The Interaction of Need, Set, and Stimulus Structure in a Cognitive Task," American Journal of Psychology, 65 (1952), 196-217.

[6]H. Helson, "Adaptation-Level as a Basis for Quantitative Theory of Frames of Reference," Psychological Review, 55 (1948), 297-313.

[7]J. S. Bruner, "A Decade of 'New Look' Perception: A Symposium," American Psychology, 14 (1959), 379.

[8]O. Hobard Mowrer, Learning Theory and the Symbolic Process (New York: John Wiley and Sons, Inc., 1960), 473 pp.

[9]C. E. Osgood, Method and Theory in Experimental Psychology (New York: Oxford University Press, 1960), 800 pp.

with language and symbolic processes. If the nature of current interest in learning is any criteria, it is difficult to find an area of research of greater concern than the focus of the present research project.

A final justification for the delimitation of the perceptive act in skill learning is that in a sense, the normative and aspirational models are the products and consolidations of the first two aspects of perception and thus form critical phenomena for evaluating the consequences of the learning in symbol differentiation and meaning.

It is agreed, therefore, that the examination of the behavioral and physiological correlates associated with the development by children of normative and aspirational models in skill learning has the potentiality for providing important and fundamental knowledge about how more effective skill teaching can take place in school.

The Nature and Limits of the Research Task

The handwriting of right handed elementary school children forms the behavioral phenomena of this study of perception: more specifically, the writing of "The quick brown fox jumps over the lazy dog."

Handwriting is a common skill task for all children receiving any kind of formal education. It thus forms a meaningful task for all children in school and is one of great personal, educational, and social significance.

The normative and aspirational perceptions of children of their handwriting thus define the problem area in this study. It was observed early in our studies of handwriting that many individuals developed a normative percept of their own handwriting which limited their handwriting behavior, i.e., such individuals would not increase the speed of their writing beyond a point where it would destroy the way they wanted their handwriting to look. On the other hand, other individuals had not developed this normative set and would thus

allow severe deterioration to take place in the quality of their handwriting under speeded writing conditions. This normative perception seemed to be the perceptual structure which allowed a child 1) to recognize his own handwriting, and 2) to value and control the quality level of his own handwriting.

The operational research tasks used, therefore, to explore this kind of perception in the three groups of children differing significantly in level of intelligence had to do with recognizing, ordering, and producing over intervals of time samples of handwriting "like mine" and "my usual" model of handwriting.

The second kind of perception, that of a handwriting model "aspired to" or handwriting "which I would want my handwriting to look like," formed the second major perceptual task of this study. The first had to do with the recognition of "how my handwriting actually is"--the normative set of "what is." The second was "what it should be." It is assumed that these two perceptions are different aspects of the perceptual process and form important parts of the total valuing act. It is assumed also that the process of learning a skill like handwriting develops a normative set, a total configuration which forms the percept "this is my handwriting," as well as an aspirational set of what the subject would like his handwriting to look like. The relationships and dynamics of these two perceptions form the basic nature of the incentives or motivational power behind skill development in handwriting.

It is recognized that there is great difficulty in contrasting one study of perception with another because of the difference in task definition. (All research workers know the critical influence of the experimental task in determining what is being studied.) Many studies have used discrimination of large numbers of pairs of objects differing multidimensionally in the formation of learning sets;[10,11,12] unidimensional problems of color,[13] size,[14] form,[15] position

---

[10] H. W. Stevenson and J. D. Swartz, "Learning Set in Children as a Function of Intellectual Level," Journal of Comparative Physiological Psychology, 51 (1958), 755-757.

with reversals and size with reversals;[16,17] and oddity and middle size tasks.[18]

Only one recent study of handwriting as a perception task is reported[19] in which a study is made of the individual's use of the beginning strokes found in words beginning with "u," "v," "w," and "y." Epstein and Hartford[20] introduced the term "secondary beginning stroke" to designate this initial upstroke to these letters since it is superfluous for purposes of legibility and tends to be eliminated with age and writing experience.

In some studies the same task is given to both normal and subnormal subjects; in the other studies reviewed, only one group of subjects--normal or

---

[11] B. J. House and D. Zeaman, "Reward and Nonreward in the Discrimination Learning of Imbeciles," Journal of Comparative Physiological Psychology, 51 (1958), 614-618.

[12] M. E. Kaufman and W. M. Peterson, "Acquisition of a Learning Set by Normal and Mentally Retarded Children," Journal of Comparative Physiological Psychology, 51 (1958), 619-621.

[13] M. B. Koch and D. R. Meyer, "A Relationship of Mental Age to Learning Set Function in the Pre-School Child," Journal of Comparative Physiological Psychology, 52 (1959), 387-389.

[14] H. W. Stevenson and I. Iscoe, "Transposition in the Feeble-Minded," Journal of Experimental Psychology, 49 (1955), 11-15.

[15] N. R. Ellis and W. Sloan, "Oddity Learning as a Function of Mental Age," Journal of Comparative Physiological Psychology, 52 (1959), 228-230.

[16] B. J. House and D. Zeaman, "A Comparison of Discrimination Learning in Normal and Mentally Defective Children," Child Development, 29 (1958), 411-416.

[17] N. O'Connor and B. Hermelion, "Discrimination and Reversal Learning in Imbeciles," Journal of Abnormal and Social Psychology, 59 (1959), 409-413.

[18] W. E. Martin and A. H. Blum, "Concept Formation and Transposition in Young Mentally Retarded and Normal Children." Mimeographed report, Contract SAE-7150, U. S. Office of Education, Department of Health, Education and Welfare, 1960, 43 pp.

[19] H. B. Linton, L. Epstein, and H. Hartford, "Personality and Perceptual Correlates of Secondary Beginning Strokes in Handwriting," Perceptual and Motor Skills, 12 (1961), 271-281.

[20] L. Epstein and H. Hartford, "Some Relationships of Beginning Strokes in Handwriting to the Human Figure Drawing Test," Perceptual and Motor Skills, 9 (1959), 55-62.

subnormal--is used. In the present study, the same writing and perceptual tasks were given to three groups of children representing different levels of intelligence, thus permitting greater comparisons to be made.

The Educational Period of Development

The educational development period covered by this study is that of the fourth, fifth, and sixth grades of the elementary school.

Since the teaching of handwriting is started in the first grade and most, if not all, of the transition to the cursive form of handwriting has been accomplished by the fourth grade, this period represents a time for elementary school children when the act of handwriting has been developed beyond the drawing stage, discriminations between and among letters and words are far advanced, writing is used to express ideas, characteristic handwriting styles have developed, and the development of normative and aspirational perceptions should be at the peak of their formation. Indeed, sixth grade children as a group have better quality of handwriting--legibility--than do adults.

It was assumed also, though unfortunately not true for fourth grade children, that children in the lower or dull group would have had educational experiences in handwriting relatively comparable to the normal and bright groups by this period in their schooling.

The period of development selected for this research, therefore, seemed to provide maximum possibilities for the discovery of differences in the perceptual behavior of the three experimental groups in relation to the assigned tasks of handwriting.

Nature of the Psychological and Physiological Correlates

Changes in the affective state of children as measured by the GSR, muscular tension as measured by muscle action-potentials, and fine motor control as measured by point pressure variability were examined in bright, dull, and normal

children as the writing conditions were systematically varied. These changes in motor and physiological states were also related to legibility of the handwriting samples produced.

Bruner and Postman[21] suggested tension and tension release as organizing factors in perception; Wapner and Werner[22] and Freeman[23] have developed the importance of motor and sensory-tonic events in the perceptual act; and Jenkins[24] has reviewed the importance of the affective processes in perception. An act like handwriting must involve a complex motor act, with definite postural and kinesthetic orientation of the human organism, and definite affective and tension states in the act of writing. The normative and aspirational perceptions of writing all involve value referents which have to do with form, structure, and levels of legibility. It seems not only desirable but necessary to study these defined correlates as a part of this study of perception in skill learning, notwithstanding difficulties in the measurement and examination of these relationships.

The Variable of Sex

In the previous studies noted, and in studies of normal and subnormal children previously mentioned, no sex differences have been identified. Differences have generally been attributed to differences in age, in background deprivation, and in diagnostic classification among the mentally subnormal.

Since work in handwriting has long noted a sex difference in favor of girls in quality of writing produced as have other studies of fine discrimination in

---

[21] J. S. Bruner and L. Postman, "Tension and Tension-Release as Organizing Factors in Perception," Journal of Personality, 15 (1947), 300-308.

[22] S. Wapner and H. Werner, "Effect of Body Status on the Kinaesthetic Perception of Verticality," Journal of Experimental Psychology, 44 (1952), 126-131.

[23] G. L. Freeman, The Energics of Human Behavior (Ithaca: Cornell University Press, 1948).

[24] H. Jenkins, "Affective Processes in Perception," Psychological Bulletin, 54 (1957), 100-127.

motor control, and there seems to be some evidence that girls and women tend to value their handwriting more than men, it seemed necessary to make sure the factor of sex could be analyzed in the design of this research. All groups, therefore, had an equal number of boys and girls and all were exposed to the same experimental conditions.

Differentiation as a Measure of Significance

The major research hypothesis of this study is that there are no differences in the three groups of children studied in respect to the perception phenomena, legibility, and motor-physiological correlates of the handwriting act. While this hypothesis permits useful analyses to be made, the decision that a given difference does or does not exist between the performance of three groups of children in respect to the legibility of their handwriting does not give support to the conclusion that intelligence is an independent variable whose variation will or will not produce corresponding variation in the legibility of the handwriting of appropriate subjects.

It should be recognized that there are two classes of variables being studied: 1) those which may be called condition variables such as intelligence and sex and 2) experimental variables such as normative and aspirational perceptual sets, point pressure, muscular tension, and speed of handwriting. In the second class of variables, only the speed of writing and the nature of writing task are systematically manipulated as the independent variables. Legibility, pressure, tension, and perceptual sets are the dependent variables. These variable relationships are studied within the conditions of intelligence and sex.

The conclusion of this study will therefore be in the nature of location of research areas of promise and the identification of further limits within which to work. Difference conclusions are not causal or even relational conclusions. They are very helpful in narrowing the focus of the research down to the variable relationships which actually control and predict the learning desired. The

results of this study should move us far in this direction, but further studies will be needed which will systematically vary the important independent-dependent variable relationships of high-low normative and aspirational set in relation to the quality of the handwriting act.

The Sampling Theory

No study reviewed which compared normal and abnormal children pretends to deal with the problems of an adequate sample in two respects: 1) in respect to an adequate sample of the full range of the intelligence continuum of elementary school children, and 2) in respect to the ability of the research population used to represent children of these characteristics generally and thus provide a sufficient base for adequate generalization.

The present study attempts to deal with a bright end of the intelligence continuum as well as with the normal and dull phases of it. Most studies reviewed earlier limit themselves to subnormal groups. The dull or subnormal group of the present study was still educable (average I.Q. 71) and was limited to a familial type of mental retardation. The intellectual distance between each group as determined by intelligence tests was approximately 29 I.Q. units--a distance so great that no errors of intelligence test measurement were likely to place any child in our sample in another intelligence group.

Effort was made to get representativeness of each group for the school population of Madison and Milwaukee, Wisconsin, metropolitan centers of 110,000 and 400,000 respectively. There are some assurances, therefore, that the conclusions of this study would generalize to other similar populations. The sampling design, while not completely adequate because of inability to find dull children who wrote with a cursive style of handwriting and who could complete the writing tasks, is still superior to any known study of its kind and purpose.

Subjects

The specific details of each research population will be described in the chapter dealing with the measurement of legibility (Chapter II) and the chapter dealing with perception (Chapter III). In general, however, two major research populations were used. The first was a Madison population of 10 individuals (5 boys and 5 girls) in the dull group, 10 individuals (5 boys and 5 girls) in the normal group, and 10 individuals (5 boys and 5 girls) in the bright group. Children nine years of age were used in 1957 to study perception in handwriting and its motor and physiological correlates; this same population was restudied two years later in 1959. The second population identified in Milwaukee dealt more adequately with the perception problem. The population of this study was carefully drawn. While a total of 60 individuals was selected for each of the three intelligence groups, groups of 30 (15 boys and 15 girls) were identified for the cross-sex comparisons for each of the three intelligence levels.

In addition, populations of writers and judges were drawn for the purpose of developing adequate legibility scaling procedures and to measure this quality in the writing samples of subjects of the Madison and Milwaukee populations.

A third sample was drawn from a summer laboratory population in 1958 to run another test on procedure and some phases of the study. This group was never used as a basic part of the research project. This sample, again, was composed of 10 children (5 boys and 5 girls) in each of the three intelligence groups.

Experimental Procedure

The details of the experimental procedures will be carefully described for each phase of the complete study in the appropriate chapter. It is worthwhile to note here, however, that all testing was done on an individual basis by trained personnel. Much of the laboratory work was done by the major investigators themselves. Every attempt was made to assure comparability of testing procedure and

to secure valid, accurate data. Repeated validity and reliability checks were run wherever possible.

Major Hypothesis

The major hypothesis of the study is that there is no difference in performance under the determined experimental conditions of the three intelligence groups in legibility of handwriting, in perception, and in motor and physiological correlates examined.

The critical questions which are used to examine this major hypothesis are detailed in each of the chapters devoted to each essential portion of the study.

General Plan of Report

This research monograph is divided into seven major sections, or chapters:

1. Statement and limitations of problem and procedures.
2. Measurement of legibility.
3. Perception of handwriting symbols.
4. Physiological and motor correlates of handwriting legibility.
5. Motor coordination and legibility.
6. Summary and discussion.
7. Appendices (containing specific details and instructions for experimental procedures).

CHAPTER II

THE EVALUATION OF LEGIBILITY IN HANDWRITING

The different tasks of perception in handwriting posed for the subjects of this study depend for their evaluation on the availability of a scale which would have the following characteristics:

1. Have the power to discriminate in legibility between several handwriting samples written by the same person under different conditions of writing.

2. Have the power to quantify differences in legibility in terms of a cardinal number system. In a cardinal scale, 3 is different than 2; 3 represents a higher value than 2; and 3 is as much different from 2 as 4 is different from 3. A cardinal scale has more power, therefore, than an ordinal scale which can only say that 3 is different from 2 and that 3 lies in a certain direction from 2. In an ordinal scale, one cannot say that 3 is as much more than 2 as 4 is more than 3.

3. Have the power to match any given sample of handwriting for the writing characteristics of slant and size of letters. It has been established that individuals can make more reliable judgments about the placement of their own writing sample on a scale when the scale sample "looks like my own."

4. Have the power to cover the total range of legibility found in the writing of middle grade children. Differences in age and educational maturity should not enable our subjects to produce handwriting samples which are either too "bad" or too "good" to be evaluated by the devised scale.

Some sense of the magnitude of the task of building a legibility scale for handwriting with these above characteristics is indicated by the following

historical review of scale building in handwriting by both previous investigators and by this Committee. It is sufficient to say at this point that prior to this study a handwriting scale with the above characteristics did not exist and that such a scale would either have to be created for our purposes or the present research problem would have to be resolved on much less adequate terms.

The Development of Scales for Evaluating Quality in Handwriting

March, 1910, and E. L. Thorndike's first report[1] on his production of a children's and an adult woman's handwriting scale in the Teachers College Record for that month marks the beginning of the development of handwriting scales in America and in the modern sense, in the world. In many ways his scaling procedures and the units into which his scale were divided were superior for research purposes to many of the handwriting scales in use today.

In rapid succession, however, the contributions of Ayres[2] in 1912 and Freeman[3] in 1915 appeared and with the work of Thorndike determined the general nature of scales for measuring legibility in handwriting for the next half century. It is true that the work of C. T. Gray,[4] Manuel,[5] Kelley,[6] Johnson and

---

[1]E. L. Thorndike, "Handwriting," Teachers College Record, XI (March, 1910), 83-175.

[2]L. P. Ayres, A Scale for Measuring the Quality of Handwriting of School Children (New York: Russell Sage Foundation Division of Educational Bulletins, No. 113, 1912). Reprinted with Gettysburg Scale, 1915.

[3]F. N. Freeman, "An Analytical Scale for Judging Handwriting," Elementary School Journal, XV (1915), 432-41.

[4]C. T. Gray, A Score Card for the Measurement of Handwriting (Austin, Texas: Bulletin of the University of Texas, No. 37, July, 1915).

[5]H. T. Manuel, "Use of an Objective Scale for Grading Handwriting," Elementary School Journal, XV (January, 1915), 269-78.

[6]T. L. Kelley, "Measuring the Quality of Handwriting," Journal of Educational Psychology, VI (December, 1915), 302-15.

Stone,[7] Gilchrist,[8] and Courtis[9] appeared in the period from 1915 to 1921, but these efforts were duplications or extensions of the important work done by these three men. It is true, also, that Starch published in 1913[10] an article on the problem of measurement in handwriting and reported his writing scale in 1919.[11,12] The Starch scales, however, were never used extensively.

Little work in scale development in handwriting was done from this period on until the revisions by West[13] and Freeman[14] of their older scales were reported in 1957 and 1959.

Outside of the work in India[15] and Sweden,[16] Beck[17] published the only new scale in handwriting developed in the past thirty years.

---

[7]G. L. Johnson and C. R. Stone, "Measuring the Quality of Handwriting," Elementary School Journal, XVI (February, 1916), 302-15.

[8]E. P. Gilchrist, "Handwriting Scale," School and Society, XII (September, 1920), 203-04.

[9]S. A. Courtis and Lena Shaw, Courtis Standard Practice Tests in Handwriting (Yonkers, New York: World Book Company, 1921).

[10]D. Starch, "The Measurement of Handwriting," Journal of Educational Psychology, IV (1913), 445-64.

[11]D. Starch, "Scale for Measuring Handwriting," School and Society, IX (February, 1919), 154-58 and 184-88.

[12]D. Starch, "Revision of the Starch Writing Scale," School and Society, X (October, 1919), 498-99.

[13]P. V. West, Manual for the American Handwriting Scale (New York: The A. N. Palmer Company, 1957). Also:

P. V. West, Handwriting: Elements of Diagnosis and Judgment of Handwriting, (Bloomington, Illinois: Public School Publishing Company, 1926). Also:

P. V. West, Chart for Diagnosing Elements in Handwriting (Bloomington, Illinois: Public School Publishing Co., 1926).

[14]F. N. Freeman, "New Handwriting," Elementary School Journal, LIX (January, 1959), 218-21.

[15]T. R. Sharma, "Construction of a Hindi Writing Scale for Primary School Children," Education and Psychology (Delhi), IV (1957), 24-31.

[16]E. Gertz, An Investigation on Speed and Quality Standards for Ability in Handwriting (Sweden: Acta University Lund, 1934).

A summary of the characteristics of the Thorndike (1910), Ayres (1912-1915), Freeman (1915-1959), and West (1926-1957) Scales is given in Table 1. An examination of these scales and the review of the above literature provide the basis for the following positions taken regarding the required characteristics of a new scale.

1. Because of the gross differences in slant and general appearance of writing, there is real question whether right and left handed writing should be mixed in a single handwriting scale. Previous scale makers argue that the general population of adults and children includes both left and right handed writers and therefore, a general scale of handwriting should sample the total writing characteristics of this population. Since the judgments made about general "merit" of handwriting are more consistent, however, when made only on the writing of right or left handed writers, we therefore felt it necessary for our research purposes to restrict our scales to right handed writing.

2. There is need to develop one master scale for evaluating and contrasting the writing samples of children of different ages and of different grades. The writing scales of Freeman and West for each grade are not usable for this reason. The Thorndike and Ayres Scales are composite scales covering the middle grades but are inadequate for our purpose for other reasons.

3. Outside of the Thorndike Scale no evidence is available for knowing the reliability of the scales when dealing with known levels of difference in legibility. When a scale has only three to five gross levels of handwriting quality, the consistent placement of a writing sample at the same level of quality is relatively easy due to the large differences in quality represented by three to five levels. If, however, the scale has 20 or more levels of possible quality difference, the reliability index is likely to be much lower than a five level scale

---

[17] M. L. Beck, "The Construction and Validation of an Analytical Handwriting Scale for Grades 4, 5, and 6," (unpublished doctoral dissertation, Boston University, 1957).

TABLE 1

CHARACTERISTICS OF SCALES FOR EVALUATION OF LEGIBILITY IN HANDWRITING

Characteristics	Thorndike (1910)	Ayres (1912,1915)	Freeman (1912)	Freeman (1959)	West* (1957)
1. Sex.	Both male and female.	Both male and female.	Both male and female.	Both male and female.	Both male and female.
2. Handedness.	Both R and L.	Both R and L.	Both R and L.	Both R and L.	Both R and L.
3. Grades included	5-8	4-6	2-8	1-8,9	2-8
4. Composite or individual grade scale.	Composite (cursive).	Composite (cursive).	Charts cover grades 2-8 (cursive).	Grades 2,3,&4, 5&6, 7&8 (cursive).	Individual grade scales on one sheet (cursive).
5. Reliability.	P.E. of individual placement,1.4 to .45.	Not given.	Not given.	Not given.	Not given.
6. Speed.	"Usual."	"Usual."	"Usual" for writing well.	"Usual" for writing well.	"Usual" but with quality.
7. Norms for speed.	None.	None.	2 minute sample, number of words per minute.	2 minute sample, number of words per minute.	Scaled for speed and quality. 2 minute sample.
8. Material copied or memorized.	Copy book needed.	Memorized and copy available.	Memorized.	Memorized.	Memorized and copy.
9. Nature of material.	A single determined sample: "Then the carelessly dressed gentleman stepped lightly..."	A single sample: "Four score and seven years ago..."	"The quick brown fox jumps over the lazy dog."	Different specimen for each grade level. Material judged appropriate for grade level.	Different specimen for each grade, familiar words, materials appropriate for level.

TABLE 1--CONTINUED

10. Criterion of placement.	Median placement by 30-40 judges.	Each sample read by 10 judges (words per minute--median words per minute is score for each sample).	Judges determine samples in three levels of excellence.	Median placement by judges.	Quality and speed.
11. Characteristics considered.	General merit.	Readability--samples of vertical, medium, and extreme slant.	Excellence in uniformity of slant, uniformity of alignment of letters, and spacing between letters and words.	Legibility, form, ease, and fluency.	Quality and speed.
12. Number of quality levels considered.	17 levels of judged equal steps. One artificially produced very bad sample rated at 4. Two 4th grade samples rated as 5,6. One copy book sample rated as 18 (best).	8 scale points obtained by assuming normal distribution. Base line from 0 to 100, median at 50 (20,30,40, 50,60,70,80,90).	5 separate charts with 3 grades of excellence. No cross grade comparisons. Summation of ratings for each of specific qualities used for 1st rating. Equal weight for all qualities but letter formation which is doubled.	Each grade scale has five levels of quality.	7 scales, one for each grade with seven levels of quality--poorest sample with slowest speed to best sample of fastest speed.
13. Nature of scale.	Cardinal.	Cardinal.	Ordinal.	Ordinal.	Ordinal.
14. Size of sample.	1000 samples, 30-40 judges.	1578 samples, 10 judges. 15780 readings, 10 for each sample.	Not given.	135,491 samples from 162 cities, 43 states; 10,646 to 22,374 samples at each grade.	Represented all parts of U.S. 644+ samples at each grade level. 42,000 judgments.

*Similar to 1927 scale.

PRECEDING PAGE BLANK-NOT FILMED

at both the good and poor ends of the continuum of writing behavior. Another study[18] suggests that this criterion discriminates among poor samples of handwriting but not among samples at the good and excellent end of a scale. Reader judges evidently do not increase (vary) their reading speed significantly when reading reasonably good handwriting, even though clear differences in "general merit" in the handwriting can be determined.

Freeman (1915) considered general excellence as a sum of the specific evaluation of the five factors of uniformity of slant, uniformity of alignment of letters, quality of line, formation of letters, and spacing between letters and words. He later (1959) used a judgment of general excellence and did not consider the evaluation of specific factors.

The problem of getting a satisfactory judgment of quality is of two kinds: 1) the inability of a judge to keep a single factor clear when evaluating a complex sample of handwriting, and 2) when single "objective" measures are made of these factors they fail to distinguish between handwriting samples which are clearly different in general merit. This seems to argue for the value of a "total" judgment of the general merit-excellence <u>or</u> quality of a sample of handwriting criterion as a basic referent for both producing handwriting scales and for placing handwriting samples on them.

West included both the criterion of speed with a criterion of quality in determining his scale, running from the "poorest handwriting produced in the slowest time to the best handwriting produced in the fastest time." It is true that speed of handwriting is associated with legibility or general quality of handwriting and needs to be considered in some fashion. Unfortunately for West's criterion, the relation between speed and legibility in handwriting is not direct but inverse and bipolar. In general, as speed increases beyond a certain point,

---

[18] Adrienne Erlebacher and Virgil E. Herrick, "Handwriting Then and Now," to be published in <u>Elementary School Journal</u>, 1961.

legibility or quality decreases. It is true that there is a range of speed within which an individual produces his best handwriting. It is just as true, however, that there are ranges of speed at both ends of his effective speed range where he produces his poorest handwriting. A scale, therefore, which assumes a straight linear relationship between increasing speed and increasing quality of handwriting is ignoring the above evidence of human learning in this skill area.

This appraisal of the different kinds of criteria used for evaluating handwriting suggests the value of a general criteria of legibility or quality as being the most valid, consistent, and differentiating basis for building a composite single scale for evaluating handwriting samples.

7. The scaling procedures of all scales, while adequate for the various purposes for which the scales were built--i.e., the evaluation of children's handwriting for instructional purposes--are inadequate for research purposes where fine quantitative discriminations need to be made on a single comprehensive continuum of writing performance. The scaling procedures used in these old scales are far less powerful than procedures available today. It is still interesting to note, however, how well Thorndike's procedure of 1910 holds up under today's examination.

8. Much was made of sampling procedures in the Ayres, Freeman, and West Scales as a basis for obtaining representative examples of the school writing of children in the various grades. Freeman, for example, reported a total sample of 135,471 individual handwriting samples collected from 162 cities over 10,000 population from 43 states with individual grade collections of 10,646 to 22,374 handwriting items. West reported 644+ representative samples for each grade for his scale. Thorndike used 1000 samples and Ayres used 1578 for their composite scales.

These numbers sound impressive but the important factor here is not large numbers but some assurance that the total nature of the range of the handwriting

quality of a given population of children has been explored. Large numbers of handwriting samples per se are not necessary to accurately describe a continuum of handwriting quality. Cities over 10,000 in America, for example, may not produce representative handwriting for all of America's children. Moreover, if a few such cities were sampled adequately, the samples from the other 150+ cities of this size would not add much more information regarding the nature and variation of legibility in children's writing.

A second factor was stressed in the scaling procedure, i.e., the number of judgments made in producing the scale (Thorndike 30,000-40,000, Ayres 15,780, and West 42,000). It should be clear that two distributions are important in scaling--the distribution of legibility in the handwriting samples considered and the distribution of judgments being made about these samples. Again, the gross size of these numbers are not valid indicators of the power of the choices being made. If the scale has only three levels of quality to differentiate the total range of writing samples, this choice can be made reliably with relatively few judges. If, however, many fine distinctions are being made in levels of quality, then the number of judgments and judges needs to increase accordingly. For the kind of judgments being made in the Freeman and West Scales, relatively few samples and judges would have been sufficient.

The Development of the Wisconsin Scales For Evaluating Quality in Handwriting

The review of previous attempts to evaluate the quality of handwriting samples revealed that no available scale was adequate for our research purposes and that new scales would have to be developed.

The Wisconsin group made five attempts to develop an adequate research scale for evaluating general legibility or quality in handwriting. These attempts in order of their chronology are:

DATE	SCALE
1952-1955	1. Six individual ordinal scales for grades 4, 6, 8, 10, 12, and college.
1958	2. A composite ordinal scale for grades 4 - 10.
1959	3. A composite ordinal scale for grades 4 - 10 with variations in size and slant for each legibility level.
1960	4. A master cardinal scale of 698 items for sixth grade handwriting measured for slant and size.
1961	5. A master cardinal scale of 600 items for fourth, fifth, and sixth grade handwriting measured for size and slant.

Initial Rationale of the Study

The general conceptions behind the various scaling projects are as follows:

1. Quality in a handwriting sample is defined as its general legibility or readability, not as its general "artisticness" or pleasing appearance.

This definition is made from the point of view of the consumer of handwriting, not from the point of view of its producer. Thus, the adequacy of a sample of handwriting to be efficiently produced, to reflect a writer's personality, or to correspond to the popular writing style of the moment is not considered. The instructions to raters, therefore, were in terms of "the ease with which you can read a given sentence." These instructions were consistent throughout all scaling projects.

While Thorndike used a judgment of general goodness or merit (a composite of legibility, beauty, and character), Ayres, Freeman, West, and Starch all agreed that "legibility" should be the most important consideration in determining quality in handwriting. This is the point of view taken by the Committee.

2. The judgment of quality should be a single general assessment of the total sample, not one based on particular aspects of qualities of handwriting within a sample. While quality of letter formation, spacing, and uniformity of

slant have all been mentioned as important factors in determining legibility, no single factor or known combinations have been demonstrated as reliably differentiating degrees of quality in handwriting.

Instructions to raters cautioned against operating on a single bias such as slant, size, letter alignment, letter formation, and density of line. They were asked to disregard neatness, spelling, and word substitutions in making their judgments.

3. The judgment of quality (legibility) in handwriting is a relative, not an absolute, judgment. It is assumed that individuals can judge Sample A and Sample B as either being alike or different and if different, can distinguish A as being better or worse than B.

With the above assumption in mind, it is recognized that:

a. There is no real zero or perfect sample in a scale derived from these judgments.

b. The number of quality levels possible in a scale is determined by the number of differences in quality reliably maintained by an adequate number of judges.

c. Scales produced on this basis are ordinal and not cardinal in character. (Level 1 is better or worse than level 3; it is not, however, three times worse or better.)

d. Scales produced on this basis are meaningful only within the population of samples which are used to produce the scale levels. A sixth grade scale produced on one sixth grade population of writing samples need not be similar either in range or level of quality with another sixth grade scale even though both distinguish six levels of handwriting quality.

The above point of view was used to develop scales in scaling projects one through three. This rationale was not used in the last two master scale item projects.

4. The writing task of "The quick brown fox jumps over the lazy dog." is considered as an adequate sample of an individual's handwriting and can be used as a valid scaling item for our purpose. It forms the writing task in all our scaling projects.

This text is short enough to be obtained easily (no need for subjects to memorize or copy), permits an easy comparison with other samples, represents all the letters of the alphabet (has no difficult words to spell), and is relatively meaningful to both children and adults. This text served as a standard writing task for all the related studies of other handwriting dimensions.

Freeman used this text in his 1915 scale but all others regarded it as being too limited in scope to serve as an adequate sample of handwriting. We felt that the use of this standard sample in a variety of writing conditions was more important than having a subject write or copy an extended writing exercise under one condition.

5. The writing was done in ink, with a standard type pen, on ruled, semi-glazed 5x8 cards. This surface permitted a clear, sharp line with little ink feathering and was judged excellent for writing with ink by a paper chemist. All writing samples for all scaling projects were obtained on these cards.

6. "Normal," "fast," and "slow" samples of handwriting were obtained from each individual in a sequence which permitted the normal sample to become the stabilized writing behavior of that individual. The normal or usual writing sample was used for all subsequent scaling projects. It was considered that this exploration of the dimensions of the writing behavior of a single individual using a single but comprehensive writing task was a more adequate picture of the writing of a single individual than a two minute sample of writing gained under conditions of memory-copying and a single set of writing instructions.

7. An adequate sample of writing for the purpose of building a scale is different from an adequate sample of writing to define the writing of a given population of individuals. The need in scale building is to obtain a population of samples of writing which is adequate to cover a universe of writing, i.e., contain a sample of the poorest writing possible, a sample of the best possible writing, and a distribution of writing behavior between these extremes which

will illustrate all degrees of quality and all the possible variations in writing style. The important thing here is to get enough independent samples so that this universe of writing is properly explored.

There is little need, therefore, to try to get samples of writing from large numbers of schools from different geographic and educational regions. It is possible, for example, to have the first 300 random samples from the fourth, fifth, and sixth grades of one school in Chicago to completely explore the writing universe which could describe writing of the 1100 elementary schools and the 186,000 children. The referent in scale building is the distribution of writing characteristics, not the distribution of children, schools, educational, or geographic regions.

8. All scales built were developed from samples of cursive writing produced by right handed individuals in school situations under controlled conditions.

Individualized Scales for Different Grade Levels:

Scaling Project I, 1952

A systematic survey was made in 1951-52 to examine the writing of school children and adults. Each subject wrote 12 different samples under changing conditions in which speed, width of point of writing instruments, and lined vs. unlined paper were systematically varied. All writing by subjects was supervised by members of our research group. Approximately 600 subjects distributed evenly in grades 2, 3, 4, 5, 6, 8, 10, 12, college, and adult groupings were used. Table 2 shows the distribution of individuals and writing samples.

The cursive handwriting samples from the preceding data-gathering operations were utilized as a basis for developing six cursive scales for grades 4, 6, 8, 10, 12, and college levels.

All writing samples which were smudged with ink blots were discarded and approximately 540 normal samples for each grade level were used to construct the

TABLE 2

DISTRIBUTION OF INDIVIDUALS AND WRITING SAMPLES
SCALING PROJECT I, 1952

Grade Level	Number of Individuals	Number of Writing Samples N W S
2	600	1800
3	600	1800
4	600	1800
5	600	1800
6	600	1800
8	600	1800
10	600	1800
12	600	1800
College	600	1800
Adult	600	1800
TOTAL	6000	18000

the scales. The specific instructions used for scaling are found in Appendix A. The number of levels of quality utilized depended on the ability of the judges to discriminate between different levels of quality. The individual scales tended to standardize, however, at five levels of legibility.

All judges were trained by the research staff and regular checks were made of their judging procedures. The criterion was that 10 judges had to agree on the placement of a single item at a given level. The 57 intercorrelations of placement of items by two judges ranged from .38 to .99. Only three of these correlations, however, fell below .70, whereas more than half were above .90.

Investigation of individual samples revealed that about 5 per cent of the samples were found to be "difficult to rate," varying widely in the judgments made by different raters in the scaling procedures.

The scales were checked for validity by rating all the samples (540 per age level) against the appropriate scale. For the most part, high agreement among different raters for the same writing sample was obtained. Again the "difficult to rate" samples caused most of the difficulty. Examination of these items revealed that these samples generally fell into three main categories: a)

extremely large or small in size, b) extreme variations in slant, and c) lack of uniformity in letter and word pattern.

Attempts to use these scales revealed the following:

1. The scales were adequate in scope to enable all obtained samples of each grade level to be scaled.

2. The placement of 95 per cent of the items could be accomplished with high reliability. Five per cent of them were placed over a fairly wide range of scale levels.

3. The five levels of legibility were so inclusive that while no item placed at a given scale level would be likely to be placed one level higher or lower, obvious differences in legibility between writing samples carrying the same rating could be observed.

4. The same handwriting sample when placed on a different grade scale would have a fairly good chance of earning a different scale value than the one gained in its original placement.

Composite Cursive Scale for Grades 4 - 10:

Scaling Project II, 1958

The need to compare the relative legibility of the writing of individuals of different ages on the same scale soon revealed the inadequacy of the individual grade scales to perform this task. An attempt was made, therefore, to build a composite scale incorporating the samples from the original scales for grades 4, 6, and 10.

The five levels were established and three samples were included at each legibility level to facilitate scaling.

This scale improved the evaluation of writing samples from different age or grade groups but judges placing writing samples on the scale still had great difficulty with about five per cent of the items which would fall into the three general categories of items identified in Project I.

Interviews with individuals doing the judging revealed that they felt more confident of their placement of items when the writing style of the sample was like that of the scale item. This observation checked with similar observations by Breed[19] and Starch[20] to the effect that "accuracy of the Gettysburg edition seemed to increase as the similarity increased between scale and copy content," (Breed) and that "accuracy in rating of samples by scales is improved when the general appearance of samples rated and of the samples in the scale itself are much alike." (Starch)

Composite Scale for Evaluating Handwriting
with Variations in Size and Slant:
Scaling Project III, 1959

The previous observations regarding how accuracy of scaling was increased when the scale sample and the subject's sample were similar in writing style led to the development of a new composite scale for grades 4 - 10; a 45 item scale which had five levels of quality, three different sizes of writing, and three different slants for each level of quality.

Our previous studies in legibility had indicated that general legibility was not affected by the factors of size and slant of letters until these factors were extreme in either direction. It seemed possible, therefore, to find samples which were representative of the variations in these characteristics in writing done by children and still be judged as representing the same general level of legibility.

Using the normal samples of writing previously obtained from children in the fourth, fifth, sixth, eighth, tenth, and twelfth grades and in college and

---

19F. S. Breed, "Comparative Accuracy of the Ayres Handwriting Scale, Gettysburg Edition," Elementary School Journal, XVIII (February, 1918), 458-63.

20D. Starch, "Methods in Constructing Handwriting Scales," School and Society, X (September, 1919), 328-9.

adult groups, eight trained judges began with the fourth grade samples and went through all the groups and examined them for specific size, slant, and legibility characteristics. As items meeting these characteristics were found, they were taken as a scale item and that scale position was considered filled. Other items from later samples which may have filled the position as adequately were ignored.

The specific procedures for determining size, slant, and legibility are found in Appendix B.

This scale enabled both evaluators and children to find a scale sample like a given sample rapidly and with high consistency. The reliability of the judging of individual writing samples from our experimental groups moved up to over .95. This quality of the scale to cover a wide range of writing behavior and to permit similarities of writing style to be found with each legibility level proved a valuable tool in dealing with our experimental materials in the perception study. A number of serious lacks, however, still remained. They were:

1. The five levels of quality were still so comprehensive that fine discriminations in legibility among writing samples were not possible.

2. While the five levels were discrete and permited consistent placement of writing samples, it was clear that the distances between legibility levels on the scale varied widely. The ordinal numbers assigned to legibility levels could not be used efficiently to arrive at defensible conclusions regarding null hypotheses of differences between legibility and other variables in our experimental design. Simply, while valuable for many evaluations, the composite scale lacked the power to deal with tasks we wished it to perform.

It was necessary, therefore, to move to a new approach to scaling if we were to develop a tool sufficiently powerful to determine the legibility distance which exists between an actual writing sample of a subject and the sample selected by him as representing his "aspired to" writing level. Or, to measure the

difference in legibility which exists between three samples of the same individual's handwriting written under the conditions of usual, fast, and slow.

Master Set of Scaled Items

for a Sixth Grade Population

Scaling Project IV, 1960

It was clear that a new scale had to have the following characteristics:

1. Define a comprehensive continuum of legibility so that:

   a. The writing of any child in our experimental group would fall within its limits.

   b. Any child would be able to find a writing sample in the scale "just like his" in general quality, size, and slant.

   c. The distance between any two points on this continuum could be determined.

2. Define slant and size characteristics of handwriting so that these characteristics for every sample would be known and comparable.

3. Apply a scaling procedure which would allow the development of a cardinal scale of legibility.

4. Develop a procedure of classification and use which would make the application of the scale to our research problems efficient.

The old procedures for scaling which in effect had followed the patterns developed fifty years ago were discarded. Rather than trying to obtain a scale with five to seven rough unequal levels of legibility with one to nine samples representing each level, it was proposed[21] that a whole population of writing samples be developed as a set of scaled items with known characteristics of legibility, size, and slant, and that this set constitute a master scale defining a given continuum of writing quality. From this master scale any number of subscales could be drawn for a variety of research purposes. This master scale would

---

[21]Adrienne Erlebacher, Research Assistant on this project, is responsible for suggesting the scaling method and for administering the scaling procedure.

specify for each item the consistency with which all of its components met the requirements of a given size or slant; and a scale which specified not only a certain quality of legibility, but which also specified the amount of variability involved in making quality judgments.

The Development of the Master Set of Items for the Sixth Grade.

Since many of our subjects were in the sixth grade, it was decided to apply our new scaling procedures to a sixth grade population of handwriting samples.

Handwriting Samples.

Handwriting samples were acquired from 27+ elementary and junior high schools in Wisconsin. Items came from various grade levels as shown in Table 3.

TABLE 3

GRADE LEVELS OF HANDWRITING SAMPLES

Grade	Manuscript	Cursive
Kindergarten	39	--
1	315	--
2	402	28
3	19	885
4	--	1000
5	--	930
6	--	954
7	--	425
8	--	289

Rating Items.

Two hundred and seventeen sixth grade handwriting samples were initially discarded because they were written in pencil or ink other than black or blue. Thirty-one were also discarded because there was more than one sentence per card. All other sixth grade samples, except six which were heavily erased or had misspelled words, were rated for legibility. The total number rated was 698. From these, 21 were eliminated from use as a final scale item because words had been

omitted or left out. Thus, a total of 677 sixth grade samples were finally admitted as items for the sixth grade legibility scale.

Rating - Raters.- Sixty-nine graduate students attending education classes in the 1959 Summer Session at the University of Wisconsin rated the 698 sixth grade samples. The data from five persons who completed 400 or less items were not used. The data from 57 persons were complete (they had rated all samples). Four others had completed 500 items and three persons had completed 600 items. From these data segments were unusable for various reasons: 1) S errors--two marks on one scale line, a mark missing from a scale line, or 2) E errors--given the wrong materials, doubt as to what order the materials had been given out.

Materials.

Samples.- The 698 "The quick brown fox jumps over the lazy dog."'s representing the samples accumulated from sixth grade students in Wisconsin which were written with black or blue ink were divided into packages of 100 items each (98 in the seventh pack) and the items were consecutively numbered. The first sentence in each pack had on it a little tag. Every tenth sample thereafter had attached to it the same kind of tag. Each tag indicated to the subject that he should be using for a tagged sample a scale line at the top of a new page of scale lines (see page 35.)

During the rating of the samples, a given pack of 100 items were shuffled three times. This was done to avoid a constant effect being exerted upon any sample by those samples which had gone before it. Thus, over the course of legibility rating, each pack of 100 samples had four different orders of items.

Also, not every person rated the same 100 items first, and so on. With several people doing the rating at one time, the tester simply gave a rater what was available in the way of a package of 100 samples which he had not had before.

Thirty-five "The quick brown fox jumps over the lazy dog."'s written with black or blue ink which were taken from the seventh grade samples accumulated in

Wisconsin were used as "warm up" material. Seven sets of five samples each were made up to comprise seven sets of "warm up" items.

Scale Lines.- The legibility rating of each item was made on an individual scale line. Each line of the scale was 9-1/2 inches in length. Ten such lines were evenly spaced on one sheet of 8-1/2 x 11 inch paper (see Fig. 1.) Ten such pages were packaged together with a staple making up a set of 100 scale lines. [For rating a package of handwriting samples S used one set of scale lines.] When "warm up" samples were given to S, an extra page of scale lines was also provided.

The following instructions were given to all E's who worked with the raters:

> Each person will be instructed exactly what to do and will then be given 105 samples and a set of scale lines containing 11 pages. On the front of the set there is a general information sheet which should be filled out by the rater (name and occupation). There is also a matrix which the experimenter can use for keeping track of which packages of samples the rater gets and in what order. It is most important that the rater not be given the same set of samples twice. Each time that the rater is given a new set of samples and a book of scales, the experimenter must write on the front of the book of scales which set of samples are being rated (i.e., is the last sample in the set 100, 200, 300, etc.).
>
> The instructions provide for the observation of the first two judgments by the experimenter. Be sure that the rater understands what he or she is supposed to do before letting him go on. However, do not yourself place a sample on the scale--do not give any indication to the rater of your personal choice and do not even give him a hypothetical example.
>
> When a person is finished with the first set of 100 (100 plus five for warm-up) samples, give him the option of a rest or of going on, and in the latter case, provide him with a second set of scales and a second package of 100 samples. Be sure to mark, each time, the number of the package being scaled on the front of the set of scales, if it has not already been done. Continue this process until the person is done with the seven packages of samples or is done for the day.
>
> If a person chooses to take a rest at any time, he may certainly do so. However, when he starts scaling again, after his rest, start him off with a set of five warm-up samples. If a person rests at the end of a set of scales, it is simple enough to start him off the next time with 105 samples and a scale book of

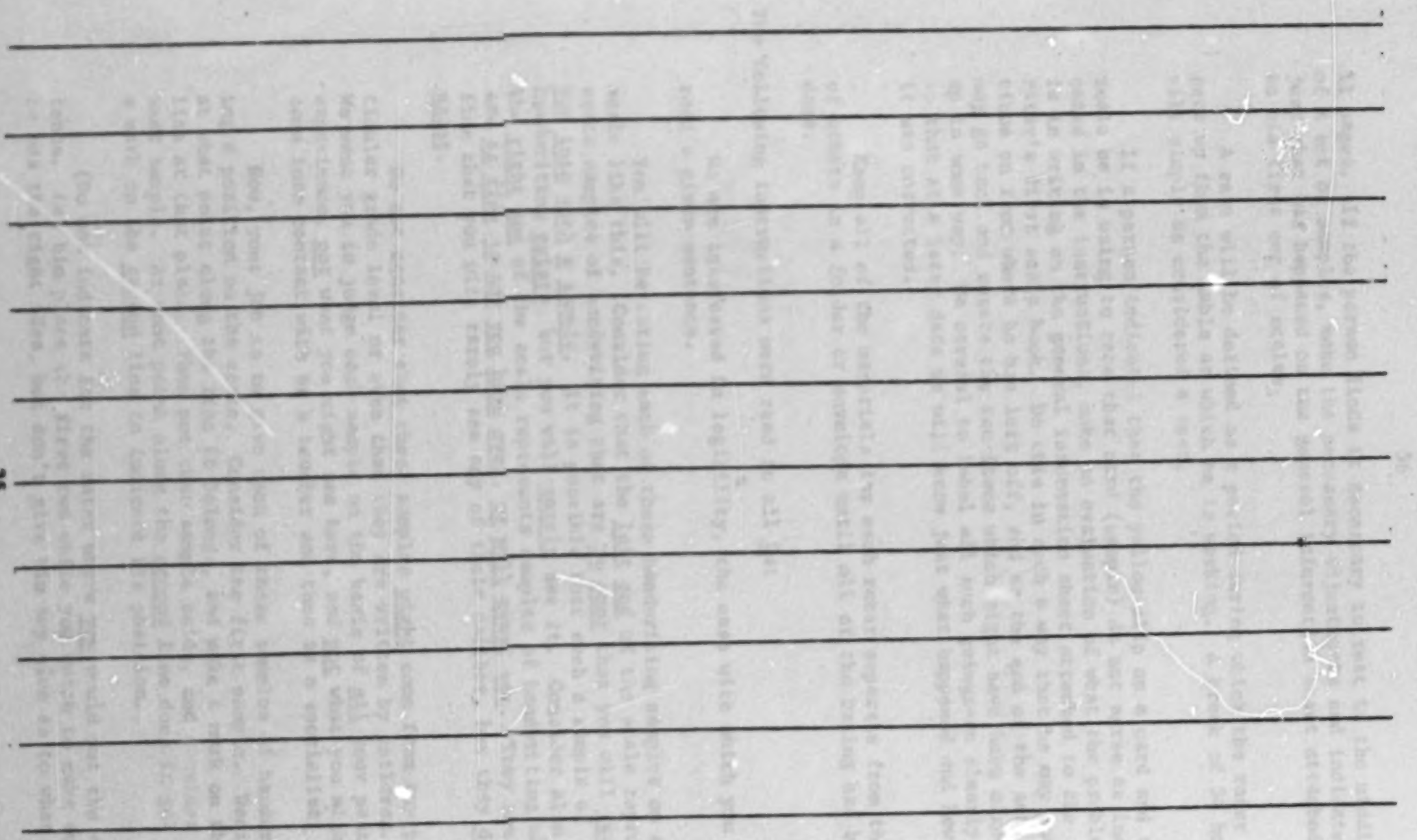


Fig. 1. Scale Lines.

11 pages. If the person finds it necessary to rest in the middle of a set of samples, make the necessary adjustments and indicate just what has happened on the general information sheet attached to his first set of scales.

A rest will be defined as a period during which the rater gets up from the table at which he is working. A break of 24 hours will simply be considered a rest.

If a person indicates that the yellow slip on a card and the scale he is using to rate that card (sample) do not agree as indicated in the instructions, make an evaluation of what the problem is in writing on the general information sheet attached to the rater's first scale book. Do this in such a way that he may continue on from where he has left off, and at the end of the set may go back and rerate the ten items which might have been mixed up in some way. Be careful to label all such goings-on clearly, so that at a later date we will know just what happened and how it was corrected.

Keep all of the materials for each rater separate from that of others in a folder or envelope until all of the rating has been done.

The following instructions were read to all Ss:

We are interested in legibility, the ease with which you can read a given sentence.

You will be rating each of these handwriting samples on a scale like this. Consider that the left end of the scale represents samples of handwriting that are so poor that you will rarely run into such a sample. It is possible that such a sample of handwriting exists, but you will rarely see it. Consider also that the right end of the scale represents samples of handwriting which are as fine as any you have ever, or will ever, see. They are so fine that you will rarely see any of their caliber, but they do exist.

Do not consider that these samples might come from any particular grade level or even that they are written by children. We want you to judge each sample on the basis of all your past experience, not what you might see here, and not what you might come into contact with as a teacher and thus as a specialist.

Now, your job is to give each of these samples of handwriting a position on the scale. Consider the first sample. Decide at what point along the line it belongs, and make a mark on the line at that place. Then put that sample aside, and consider the next sample. At what point along the second line does it go? Make a mark on the second line to indicate its position.

(Do not indicate for the rater where you would put the sentence. Let him place the first two while you watch to make sure he has the right idea, but don't give him any clue as to what you

think is the right placement. He should not be trying to please you. He should only be trying to make honest judgments.)

After you get going (after the five warm-up sentences are done), every tenth sentence will have a little yellow tag on it indicating that you should be using a scale at the top of a page. If you find that you are not "on schedule" let the person in charge know about it and he will straighten things out.

When you finish a package of cards, let the person in charge know. If you want to take a break at that time, fine; if not, the person in charge will bring you more cards.

Start at the beginning. Work quickly. We have found that we get the best results when you work quickly--accurately, but as quickly as you can comfortably go.

A transparent grid was made which divided each 9-1/2 inch scale line into 24 units designated 1 through 24, representing 24 legibility ratings.

The rating given by each of the 64 raters (a mark on the appropriate scale line) to each of the 698 items used was noted and the ratings for each item were cumulated.

Method of Scaling.- The items were then scaled by Attneaves'[22] method of graded dichotomies. This method is based essentially on Thurstone's Case V of the Law of Comparative Judgments. Here it is assumed that a rating placed at a certain position on the continuum for one writing sample is the same psychologically speaking as a rating placed at the same point for another sample. Furthermore, for each sample there is a normally distributed discriminal process with dispersions having equal σ's. Once these distributions have been identified, through the application of arbitrary cutting points, the psychological units are

---

[22] F. Attneave, "A Method of Graded Dichotomies for the Scaling of Judgments," Psychological Review, Vol. 56, 1949, pp. 334-340. Also, F. Attneave, "Dimensions of Similarity," American Journal of Psychology, Vol. 63, 1950, pp. 516-56. For additional discussions of procedures, see: Warren S. Torgerson, Theory and Methods of Scaling (New York: John Wiley and Sons, Inc., 1958). Chapter 10, "The Law of Categorical Judgment," pp. 205-245. Also, W. R. Garner and H. W. Hake, "The Amount of Information in Absolute Judgments," Psychological Review, Vol. 58, 1951, pp. 446-59, and A. L. Edwards and L. L. Thurstone, "An Internal Consistency Check for Scale Values by the Method of Successive Intervals," Psychometrika, Vol. 17, 1952, pp. 169-80.

defined and the scale values for each sample may be determined. Because of these assumptions the method requires a large number of raters and a fair number of stimulus objects to scale.

Often a single error on the part of an J made it necessary to eliminate a page of scale lines from consideration. Table 4 gives the number of ratings

TABLE 4

NUMBER OF RATINGS USED IN EVALUATING ITEMS

Number of Ratings Used		Number of Items Using Each Rating
64	.....	97
63	.....	233
62	.....	102
61	.....	119
60	.....	120
59	.....	4
45	.....	10
44	.....	20

used in evaluating each item.

The reliability of the tabulating procedure was evaluated by independently tabulating a second rating for 20 randomly selected items. The scale value given each of the 20 items the second time was compared to the rating given the first time in coordination with establishing the scale. A Pearson Product Moment correlation between original and re-scored items was .99.

Recording of Legibility Scores.- The legibility scores ranged from 1.16 to 4.91 for the 677 sample items. These scores were recorded on a master card for each sample.

The Determination of Size.

Materials and Judges.- Each sentence of the 677 scaled samples was scored for size. The scoring in each case was done by two judges acting independently.

Instructions to Judges.- The following was the set of instructions for size measurement of scale items ("base letters") given to each judge:

Use as a measureing instrument No. 340-MP Ditagen Paper, Millimeter size. With pen, mark off blocks the width of the paper 1mm., 2 mm., 3 mm., etc., through 10 mm.

Place the measuring paper over a given sample of handwriting. Place the base line of the 1 mm. block parallel with the base of the first word, "The." Fit the base of the 1 mm. block as close to the lowest "base letter" extension of that word as possible: even include ink into the base line of the 1 mm. block--but do not let any part of a base letter appear (peak out) below it. Record the number of base letters (from a possible two--the "e" and the base of the "h") which fit within (you do not have lines extending above) the 1 mm. block. Do the same for each of the following words, with each word, being careful to line up the base line of the 1 mm. block with the base of the word being counted. Record the total number of letters (from 31 possible) which fit in the 1 mm. block. If not all 31 letters are accounted for, go on to the next largest block, i.e., the 2 mm. block. Follow the same procedure which was followed before, including as fitting within the 2 mm. block, any letters that fit in the 1 mm. block. That is step one. Continue the procedure until all 31 letters are accounted for.

Then (step two), subtract appropriately to find out how many letters fit within one and only one block (i.e., if 15 letters fit within the 2 mm. block and 5 fit within the 1 mm. block you know only 10 fit within the 2 mm. block and not the 1 mm. block.)

When you are done with any sample of handwriting, you will have a set of numbers which will tell you how many letters are 1 mm. or smaller, how many are between 1 mm. and 2 mm., etc. We will know the range the handwriting has and which size it is on the average.

The letters to be considered are the following:

he quick brown ox jumps over he azy dog

The upstroke on the last letter of a word is part of that last letter. Don't count the cross on the "x."

Scoring.- Each of the samples was measured twice, independently, for size, according to the instructions and given a mean score. An example of scoring item 1 by persons "a" and "b" follows. Scores of "a" and "b" were kept separate until step 3 and are shown together here only to facilitate explanation of scoring.

Step 1
Initial Measuring

Sample	1 mm. or less	2 mm. or less	3 mm. or less	4 mm. or less	5 mm. or less	6 mm. or less
1a	0	5	26	31		
1b	0	4	24	31		

Step 2
Subtraction

Sample	1 (0-1 mm)	2 (1-2 mm)	3 (2-3 mm)	4 (3-4 mm)
1a	0	5	21	5
1b	0	4	20	7

Step 3
Total Number of Letters Put
In Each Category by "a" and "b"

Item 1	1	2	3	4
Total	0	9	41	12

Step 4
Compute the Mean Size Score

Item 1 has:	0 no. 1's	0 x 1 =	0
	9 no. 2's	9 x 2 =	18
	41 no. 3's	41 x 3 =	123
	12 no. 4's	12 x 4 =	48
			189 = total size

189 ÷ 62 = 3.05 = the mean size

The item range was also noted, i.e., the lower and upper bounds of letter size were noted, and the range (over one, or two, or three, etc. sizes) was specified.

Reliability.- A random sample of 30 items was drawn from the sixth grade item. Mean scores for measurer-1 and measurer-2 were compared with mean scores of measurer-3 and measurer-4.

A Pearson Product Moment Correlation coefficient of .94 was found. The differences between mean 1 and 2 scores and mean 3 and 4 scores ranged between 0 and .47.

The Measurement of
Slant in Writing Items.

Materials and Judges.- Each sentence of the two scales was scored for slant. The scoring in each case was done by two judges acting independently.

Instructions to Measurers.- The following was the set of instructions for slant measurement of scale items ("loop letters*") given to each measurer:

> Use as a measuring instrument the Number 2 Slant Scale (See Appendix B). The 20° of the left of ⊥ slant (backwards) is designated number 1, the perpendicular is designated number 2, the 20° slant to the right of ⊥ is designated number 3, and the 40° slant to the right is designated number 4.
>
> Lay the scale over each sentence and determine the slant of each of the 15 loop letters individually. (Do not consider just the loop; consider the whole letter.) Choose that slant which best fits the whole letter, i.e., best fits right through the middle. Record the number of loop letters that are of each slant.
>
> When you are done measuring the slant on each sentence, indicate whether or not you think the slant of the loop letters is representative of the sentence as a whole.

Scoring.- Each sentence was measured twice, independently, for slant. The two independent measurements were totaled, and the sentence was scored as follows:

Given the total number of measurements equal to thirty (30), indicate 1) the number and 2) the per cent of this total possible-score in each of the four slant categories.

The number of letters falling in any slant category was given by straightforward count. From this a mean slant was computed.

The simple coverage slant score for each item was computed as follows:

	Slant			
	#1	#2	#3	#4
Sample "x"	2	3	16	9

Two letters measure slant - 1
Three " " slant - 2
Sixteen " " slant - 3
Nine " " slant - 4

Therefore
1 x 2 = 2
2 x 3 = 6
3 x 16 = 48
4 x 9 = 36
92

$$30\overline{)92} = 3.07$$
90
200

Average slant = 3.07

*Loop letters are : t h q k b f j p t h l z y d g.

The per cent is indicated as follows and is illustrated in Table 5.

TABLE 5

INDICATION OF WEIGHT OF EACH OF THE FOUR SLANT CATEGORIES

Total Number of Measurements in a Given Category	Per Cent of Total Included	Score Notation	
0-5	0-19.99	0	0
6-11	20-39.99	-	1
12-18	40-60	=	2
19-24	60.01-80	≡	3
25-30	80.01-100	≣	4

Any category which has 0-5 measurements recorded therein (0-19.99% of the total) is given a 1 (-); any category which has 12-18 measurements recorded therein (40-60% of the total) is given a 2 (=); any category which was 19-24 measurements recorded therein (60.01-80% of the total) is given a 3 (≡); and any category which has 24-30 measurements recorded therein (80.01-100% of the total) is given a 4 (≣).

Reliability of Scoring Procedure.- The following question was asked: "If the measurement of each of two measurers were considered, would we find that the two agree as to what the measurements for any given sentence are?" To answer this question, two persons measured independently the first 100 sixth grade sentences. Then, for each of the two measurers, we found independently the score (not the mean score but the profile or weight of each category) on the basis of a total of fifteen measurements for each sentence. Measurers scored weighted categories of 41 out of the hundred pairs identically. The measurers agreed on which categories were used (i.e., which categories had in them 20% or more of the total measurements regardless of what per cent over that) in 60 out of the hundred pairs.

The question was then raised as to the reliability of the scoring (still using the weight system) done on the basis of the total scores of two measurers. The measurements of measurer-1 and measurer-2 for the first one hundred scores were available. These were totaled for each item. We had two additional independent measurements made for each of the first 100 items (making a third and fourth measurement). The third and fourth measurements were totaled for each item. Each of the 100 items was scored twice, using the measurer-1 and measurer-2 total, and the measurer-3 and measurer-4 total. The two sets of scores were then compared.

Fifty-four cases of non-identical scores (weights) were found. It was noted, however, that if the criterion for agreement was "agreement within one juxtaposition between adjacent categories of 20-40% of the scores," a very high per cent reliability could be found. This may be illustrated by:

Scores as Measured by:	Category 1	2	3	4
Measurers 1 & 2	0	0	-	■
Measurers 3 & 4	0	0	0	■

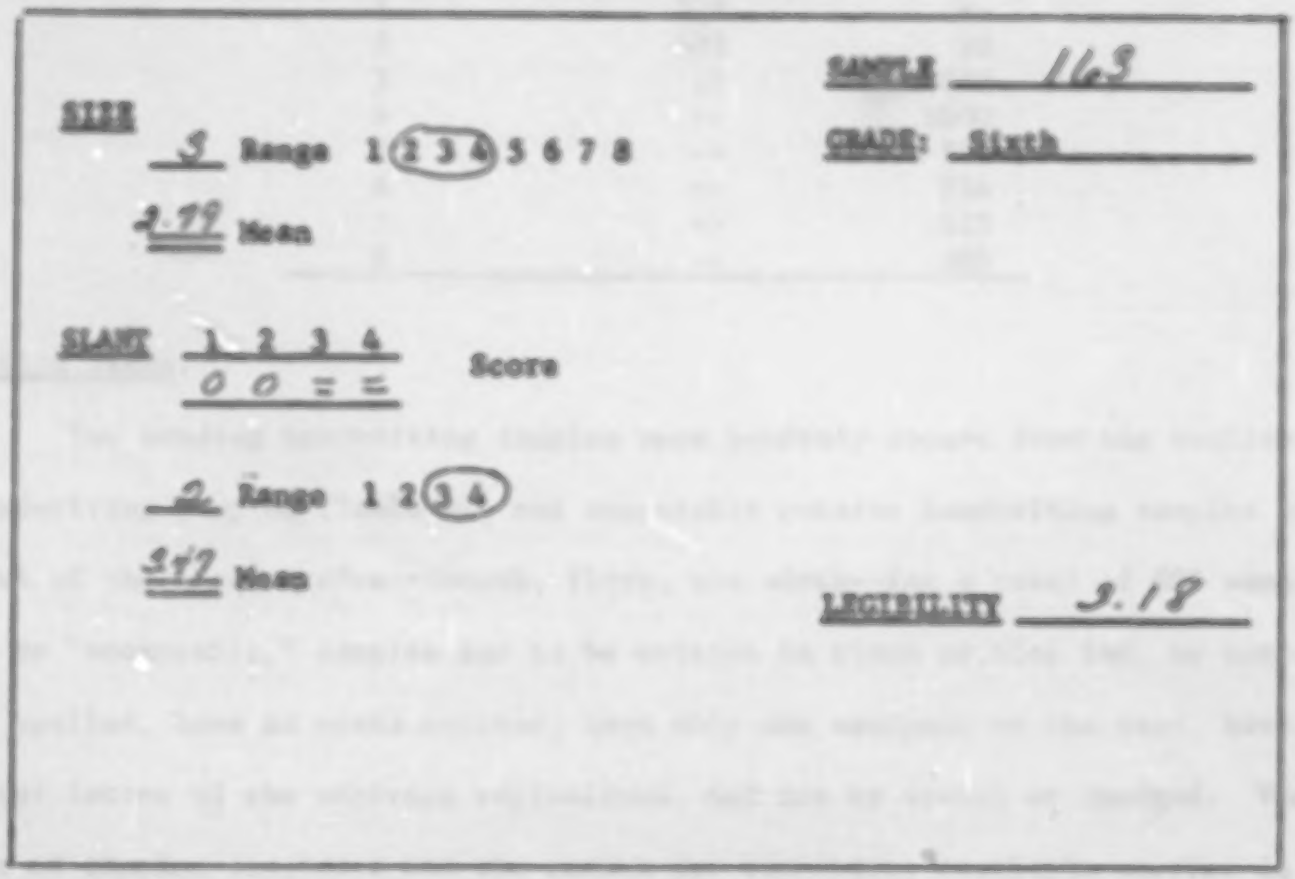
Note that there is not immediate agreement as to the scoring. However, if the "-" (20-40% of the scores) in 1 and 2's category 3 is moved to category 4, or if one of the "-'s" (20-40% of the scores) in 3 and 4's category 4 is moved to category 3, they are equivalent in score. This is what the writer calls the juxtaposition between adjacent categories of 20-40% of the scores.

Calling this "agreement," 91 out of 100 scores were found to be in "agreement." When obvious errors (clearly instances of putting the counted measurements for one category in the wrong place--the majority of the measurements not being in the same category) were corrected, there was 93 per cent agreement. It was thus decided to use judgments of two judges for scoring every item.

<u>Reliability</u>.- For purposes of computing reliability, the weighting system was abandoned and the mean slant score used.

Thirty samples were drawn at random from among sixth grade items for the purpose of computing the reliability of the composite score of two judges. The 1 and 2 mean scores were compared with the 3 and 4 mean scores to determine the mean agreement. A Pearson Product Moment Correlation coefficient of .98 was found.

Recording.- The slant and size scores were recorded on the master card. This card for item 163 is reproduced in Fig. 2.


SAMPLE 163

SIZE

5 Range 1 (2 3 4) 5 6 7 8

2.99 Mean

GRADE: Sixth

SLANT 1 2 3 4

0 0 = =

Score

2 Range 1 2 (3 4)

3.47 Mean

LEGIBILITY 3.18

Fig. 2. Master Recording Card.

Master Item Scale for Grades 4, 5, 6

Scaling Project V, 1961

The sixth grade scale proved the validity of the scaling and scoring procedures but failed to provide a comprehensive base for contrasting writing items from populations over grades 4 5, and 6. It was necessary to build this master set of items to deal with our perception data from the Madison and Milwaukee population.

Handwriting Samples.

Handwriting samples were acquired from 27+ elementary and junior high schools in Wisconsin. Items came from various grade levels as shown in Table 6.

TABLE 6

GRADE LEVELS OF HANDWRITING SAMPLES

Grade	Manuscript	Cursive
Kindergarten	39	--
1	315	--
2	402	28
3	19	885
4	--	1000
5	--	930
6	--	954
7	--	425
8	--	289

Rating Items.

Two hundred handwriting samples were randomly chosen from the available handwriting samples (Table 6), and acceptable cursive handwriting samples of each of the three grades--fourth, fifth, and sixth--for a total of 600 samples. To be "acceptable," samples had to be written in black or blue ink, be correctly spelled, have no words omitted, have only one sentence on the card, have the first letter of the sentence capitalized, and not be erased or smudged. The number of samples discarded and the reason for discarding available samples is shown in Table 7 for each grade. The number acceptable is also shown.

TABLE 7

ACCEPTABLE AND UNACCEPTABLE HANDWRITING SAMPLES

Grade	Number and Reasons For Discarding Samples			Number Acceptable
	Pencil or Colored Ink	2 or More Sentences	Other Categories	
4	550	60	60	328
5	242	0	65	625
6	217	31	6	698

Raters.

One hundred eleven graduate students attending education classes in the 1960 summer session at the University of Wisconsin rated the 600 samples from the fourth, fifth, and sixth grades.

Rating Data.

From the ratings of the 111 judges, those of 16 persons were not used because of mixups in individual packs, incorrect following of instructions, prejudice due to previous testing, leaving out a scale sheet, and the like.

The data from forty-two persons were used complete. The data from 50 were incomplete due to mismarking of the scale sheets. The data from 34 of these persons were used with ratings of 30 or 40 items missing from each in such a fashion as to make up 32 sets of complete (60 items) data. This was done to facilitate analysis of the data on the IBM 650 digital computer.

Preparation of Material.

Samples and Packs.- In order to eliminate any bias for a particular grade or school, the six hundred samples of "The quick brown fox jumps over the lazy dog." were shuffled. They were numbered consecutively. To insure that the cards stayed in numerical order each pack was bound with string.* Each page was given a code number. Thus, handwriting samples one through 10 formed a pack identified by the number one, samples 11 through 20 were pack number two, and so on. Warm up material in the form of one of several sets of 10 fourth and fifth grade handwriting samples were set up. Additional items were made available for supervised practice.

The order of presentation of the packs to raters was changed each day by using a random numbers table to determine the order (the order within each pack remained the same). So pack one was not necessarily the first one given to a

*This was not done for the first few Ss and much data was unusable.

rater but the samples within always ran consecutively from one through ten. Each subject was given all 60 packs. The packs were given out in groups of 10. There were thus six groups of 10 packs of 100 samples each. Since more than one person of 10 rated at a time, a person did not necessarily start with the first group in the day's order but started with the 11th, 21st, 31st, etc. pack for the day, i.e., with any available. Subsequent groups were given out as they were available. The entire procedure took usually less than one hour.

Scale Sheets.- The legibility rating of each item was made on an individual scale line. Each scale line was 9-1/2 inches in length. Ten such lines were evenly spaced on one sheet of 8-1/2 x 11 paper (see page 35). Each page of 10 scale lines had on it a number which corresponded to a package of 10 handwriting samples. Each pack of samples that S rated he rated on the pages of scales with the corresponding number. When a set of 10 warm up items was given, a scale page numbered 61 was used.

The scale sheet consisted of three pieces of paper of which S saw only the outside. The one which the subject saw consisted of 10 lines--three quarters of an inch apart--running parallel across the page. Underneath this was a carbon sheet which transmitted the subject's recordings onto the last paper. This last paper was exactly the same as the first (the lines of the first page were superimposed directly over the lines of the last) with the addition of a set of lines which crossed the aforementioned 10 lines perpendicularly at one centimeter intervals forming a grid pattern of twenty-four columns. The columns were numbered 1 through 24. The first and last pages were actually one sheet folded with the carbon inserted in between.

Procedure for Testing.

The subjects were asked to fill out a general information sheet with name and occupation on it and a place for E to keep track of which group of samples

had been given. They then read the set of instructions and rated a sample pack of cards on a mimeographed scale sheet. This sample scale sheet was used for practice only and was later thrown away. When a subject had finished rating the samples and all questions were answered, the experimenter verbally reinforced the instructions. Ss were asked to review the procedure for making judgments. Those ideas that received particular stress were: 1) that the subject was to rate the samples according to how quickly and easily he could read them; 2) that the subject was to work as quickly as possible; and 3) that any irregularities were to be brought to the attention of the experimenter. The subject was then given the first group of 110 items (a pack of 10 warm up items was included at the beginning of the group of items the subject received first.) The other five groups were given subsequently.

When an error was called to the experimenter's attention she simply replaced the scale sheet with a fresh one appropriately numbered and the subject rated that pack of cards over again.

Instruction to Raters.

1. We are interested in legibility, the ease with which you can read a given sentence.

2. You will be rating each of these handwriting samples on a scale like the one given you. Consider the left end of the scale representing samples of handwriting that are so poor--so illegible--that you will rarely run into such a sample. It is possible that such a sample of handwriting exists, but you will rarely see it. Consider, also, that the right end of the scale represents samples of handwriting which are as fine--as legible--as any you have ever, or will ever, see. They are so fine that you will rarely see any of their caliber, but they do exist.

3. Do not consider that these samples might come from any particular grade level or even that they are written by children. We want you to judge each sample on the basis of all your past experience, not what you might see here, and not what you might come into contact with as a teacher and thus as a specialist.

4. Now, your job is to give each of these samples of handwriting a position on the legibility scale (i.e., on one of the lines provided). Consider the first sample. Decide at what point along the line it belongs, and make a mark on the line at that

place. Then put the sample aside, and consider the next sample. At what point along the second line does it go? Make a mark on the second line to indicate its position. Continue with the rest of the sample which you have.

Procedure for Tabulation.

After the testing was completed E checked each scale sheet. This was done for two reasons: first, to clarify any ambiguity in the markings; and second, to check for errors. In the first case, when any part of a rater's mark intersected one of the lines which formed the boundary of one of the 24 cells in the grid pattern a red pencil line was drawn to indicate which column the mark fit in. The experimenter's procedure for judgment varied with the type of mark made. If, as in most cases, it was an "x" the score was judged to fall within that square in which the intersecting lines of the "x" lay. However, if a mere bar / was used, the experimenter then determined its value by its position on either one side or the other of the intersecting lines of the grid. When this had been completed the top sheet of the scale (the one on which the judge had made his mark) and the carbon paper were thrown away. Those scale sheets that contained errors, e.g., placing two marks on one line, were left complete and labeled as to the nature of the mistake. A perfect pack of sixty, all from one judge, was divided into two packs of thirty, stapled, and identified by name.

The data used were then transferred to IBM cards and scaling was completed on an IBM 650 programmed for Attneave's Method of Graded Dichotomies for the Scaling of Judgments.[23] Seventy-four judgments were used for evaluating each item.

Summary

It has been possible:

1. To develop a master set of 600 scaled items for grades 4, 5, and 6 which provides:

---

[23]Attneave, loc. cit.

a. A definition of a comprehensive continuum of writing legibility for this age school children.

b. A cardinal scale of legibility which will permit computation of known differences in legibility between two samples of writing.

c. A means for identifying scale samples of similar style through a procedure for measuring slant and size characteristics of writing.

d. A research tool which has greater power for dealing with legibility in handwriting than any scale that now exists.

2. To develop a base for the evaluation of legibility of handwriting which can be used in programs of handwriting instruction in the elementary schools of America.

CHAPTER III

CHILDREN'S PERCEPTION OF THE HANDWRITING TASK

Chapter I described the various models into which the act of perception can be divided. The normative model, the way in which the perception of the task appears to operate in setting certain controls on an individual's behavior, and the aspiration model, the way in which the individual appears to use his perception as a referent for qualitative judgments about his performance, are the two perception models of greatest interest in the present study. Little attention is paid as to how a subject came to be able to distinguish one letter or word form from another, or how meaning came to be attached to a word like "brown." As was explained, these aspects of the perception process are important but have been extensively explored. The attention here is directed towards exploring the degree to which a child has established a "normative" conception of his own handwriting so that he can maintain it, identify its place in a continuum of handwriting, and distinguish it from the "aspired to" model. The exploration of these two aspects of perception is designed to take place with children who have already learned and established their handwriting and who are divided by widely differentiating conditions of intelligence.

The Problem

The basic question posed in this research is, "Do children of different levels of intelligence perceive given handwriting tasks in the same way?" The experimental hypothesis is that they do. The operational research hypotheses are formed by the designed research writing tasks themselves and will be described later.

Procedure

Four studies[1] were designed by the Committee for Research in Handwriting to gain a more effective understanding of the role of perception in the handwriting act. The salient background data for these studies are given in Table 8.

Note that discrete levels of intelligence between groups were established, that both boys and girls were tested, and that there are certain differences in the studies in experimental conditions, in type of writing (manuscript or cursive), and in handwriting scales used. Laboratory approaches were used to explore certain phenomena (studies 1 and 3), and to provide a basis for the more extensive field study (study 4). This report is directed toward the latter of these studies (studies 3 and 4) for which more extensive and validated data on children's perception of the handwriting task is available.

Sampling Distributions for Individual Studies

Essentially, two populations were used for this project: a Madison population of bright, average, and slow learners were used for the laboratory studies 1 and 3; and a Milwaukee population which was representative of the populations of bright, average, and slow learners in the metropolitan area was used for study 4. The results of study 2 were used only as a note to elaborate certain phases of the population data because of unusable polygraph records.[2]

The Sample for Study 1 and 3

The population used for studies 1 and 3 was drawn from the total population of children attending the Madison Public Schools in 1956-57. Professor

---

[1]Valuable assistance was furnished in this phase of the study by Mr. Dan Andersen, Research Assistant, and also by Mrs. Adrienne Erlebacher, Miss Nora Okada, Research Assistants, and by Miss Leatrice Anderson, former Research Assistant.

[2]This study was planned to see how the results gained in a population of children coming from a widely varied educational and social background (Summer Laboratory School) but differentiated by intelligence on the same perceptual and motor tasks would differ from our controlled laboratory and field populations.

TABLE 8

BACKGROUND DATA FOR FOUR STUDIES OF THE HANDWRITING OF BRIGHT, AVERAGE, AND DULL I.Q. CHILDREN

Study		Date	N			S's CA (Mos.)	I.Q.		Type of Test	Experimental Conditions	Type of Writing		Scale Used In Perception Choices	Other Aspects of Handwriting Tested
			M	F	Total		Ave.	Range			C	MS		
1	B	1957	5	5	10	9-1	129	125-143	Individual and Group	Laboratory	x		Fourth Grade (mixed)	Force Variation Speed of Writing Motor Control GSR
	A		5	5	10	9-2	100	94-105			x			
	D		5	5	10	9-1	71	64-78				x		
2	B	1958	4	5	9	10-0	135	125-161	Group	Laboratory	x		Fourth Grade (mixed)	Electrical Artifacts made polygraph records invalid
	D		5	4	9	9-8	94	83-99			x			
3*	B	1959	5	5	10	11-1	129	125-143	Individual and Group	Laboratory	x		Individualized (size and slant)	Force Variation Speed of Writing Motor Control GSR Action Potentials (neck and arm)
	A		5	5	10	11-2	100	94-105			x			
	D		5	5	10	11-1	71	64-78				x		
4	B	1959	17	18	35**	10-2	134	116-171	Individual and Group	Field	x		Individualized (size and slant)	Speed of Writing Motor Control (Paper and pencil test)
	A		17	18	35**	10-1	100	90-110			x			
	D		19	16	35**	10-7	68	51-83			x	x***		

*Identical with group 1 except for the substitution of five cases judged equivalent.
**See section on explanation of sample. Thirty subjects in each category with complete cursive records.
***The 25 subjects with manuscript writing not used.

Herbert Klausmeier[3] had surveyed the Madison school population as to the number of gifted and subnormal children and had proceeded with the giving of individual Wechsler Intelligence Tests. This school population with the related background of data was shared with Professors Klausmeier and Ringness.[4]

The experimental sample consisted of 5 boys and 5 girls of low intelligence, 5 boys and 5 girls of average intelligence, and 5 boys and 5 girls of high intelligence. The criteria were I.Q., birthdate, and parental approval to participate in the study. Low intelligence was defined as an I.Q. of 64-78 (average 71) as ascertained with the Wechsler Intelligence Scale for Children;[5] average intelligence as a WISC I.Q. of 94-105 (average 100), high intelligence as a WISC I.Q. of 125-143 (average 129). The average chronological age of these groups was 9 years, 1 month (low), 9 years, 2 months (average), and 9 years, 1 month (high). All children in the project were screened for vision and hearing. Low I.Q. children were excluded who exhibited characteristics of definite organic types such as mongoloid, hydrocephalic, microcephalic, and cretin; those severely handicapped orthopedically; and those suffering from epilepsy which was not under good control with medication.

While not essential to the major design of the major study, the findings of this study would have been a useful check on the findings of the laboratory and field studies.

[3]Herbert J. Klausmeier, An Analysis of Learning Efficiency in Arithmetic of Mentally Retarded Children in Comparison with Children of Average and High Intelligence, Contract #153 (6433), Department of Health, Education, and Welfare, United States Office of Education (Madison: School of Education, University of Wisconsin, 1959).

[4]T. A. Ringness, Emotional Reactions to Learning Situations as Related to the Learning Efficiency of Mentally Retarded Children, Contract SAE 6434, Department of Health, Education, and Welfare, United States Office of Education (Madison: School of Education, University of Wisconsin, 1959).

[5]For those children in the average group who did not have WISC scores, California Tests of Mental Maturity, Kuhlman-Anderson, and individual Binet scores were used. Because of the high correlation which exists between these tests and the difference in average intelligence which exists between these groups (29 I.Q. points), it was assumed that no individual child would be placed in another experimental group by variations among the different intelligence tests used.

The individual children were chosen randomly within the criteria limits. Children in the bright group were randomly selected from all the gifted children in a given school of 810 children. There is no reason to believe that this group would be any different from any other group of randomly selected children from a group meeting these criteria. The average group was selected randomly from all the children meeting these criteria in the Madison Public Schools. The children of the dull group were selected from two schools having special classes for the mentally retarded. All of these children were in the third and fourth grades in Madison schools.

General Background Data for Sample

General Health

Using the data gathered from medical examinations given by Klausmeier, the medical records of the school, and the reports of individual reports of their teachers, it was concluded that the children were in good health with no significant differences according to health in the I.Q. groups.

Psychological Health

The Rorschach, three cards from the Thematic Apperception test, and one Draw-A-Figure test were given by a clinical psychologist in one sitting with the greater portion of the subjects in the high and low intelligence groups (Klausmeier).

The results of these tests and interviews, reports of teachers, and observations of children in an experimental setting gave support to the conclusion that all the children in the sample had sufficient mental health to attend to the conditions of the study and were able to bring their experience and abilities to bear on the testing situations. Certainly these children were representative in this respect to similar samples which might be drawn from school populations in cities of like size and socio-economic conditions.

Occupation of Father

The bright and average groups of children had fathers who commonly came from the professional and managerial occupational category while children in the lower group had fathers who were classified in the skilled and unskilled occupations. This difference was judged to be characteristic of what would be true of samples drawn on a similar basis from general school populations.

Etiology of Low I.Q. Children

Klausmeier found that the reports of the examining physician and psychologists on the low intelligence of this sample as to whether the intelligence level of the child was due to organic or hereditary factors were inconsistent, a fact which is well documented in the literature. Since the average I.Q. of the low group was 71, there was no reason to attribute this retardation to hereditary or organic causes any more than to impoverished environment.

Both Gallagher[6] and Capobianco[7] report no differences between brain injured and familial mentally retarded children on tasks involving perceptual ability and in arithmetic achievement. It was judged, therefore, that lack of further information of this nature did not constitute a serious handicap to the study.

Changes in Sample Over Two Years

The 1957 sample (third and fourth grade) was retested two years later in 1959 (fifth and sixth grade) to examine what the effect of this additional maturity and educational experience would indicate for their behavior in certain of the experimental tasks.

---

[6]James J. Gallagher, "A Comparison of Brain-Impaired and Non-Brain-Impaired Mentally Retarded Children on Several Psychological Variables," Society for Research in Child Development, 22 (1957), 1-79.

[7]Rudolph J. Capobianco, "A Comparative Study of Endogenous and Exogenous Mentally Handicapped Boys on Arithmetic Achievement," Dissertation Abstracts, 14 (1954), 794-95. Abstract of Ed.D. Thesis, University of Illinois.

Five children of the original 1957 group had left the city and were not available. Five additional children, therefore, were selected to represent these children as adequately as possible. The group characteristics remained unchanged with these replacements: Dull group C.A. 11-1, average I.Q. 71, 64-78; Average group C.A. 11-2, average I.Q. 100, 94-105; High group C.A. 11-1, average I.Q. 129, 125-143. It was judged, therefore, that no important modification had been made in the experimental groups.

Cooperation of Parents
and Transportation of Children

Superintendent Phillip Falk, Madison Public Schools, and the School Board of Madison approved the project on March 5, 1957. Dr. Carl Waller, Assistant Superintendent and Director of Child Study, and Mrs. Helen Holcombe, Director of Special Education, participated in the identification and selection of the children. Mr. Ralph Allen, Principal of the Randall School, Miss Pearl Lerous, Principal of the Washington School, and Mr. A. Kermit Frater, Principal of the Lowell School helped in securing permission of parents, the cooperation of teachers and children, and also in overseeing the logistics of getting children from their classrooms to the University Laboratory and back.

Only one parent refused our request for cooperation and would not sign the permission form. The copies of the letter and permission form are found in Appendix C.

The children were transported to the Education Building on the University campus by licensed carriers and were returned to their respective schools by the same means. These children were accompanied at all times by a research assistant or a member of the research staff. There were no unfavorable incidents during either the 1957 or 1959 testing periods. There was ample indication that children participated fully and with enthusiasm in the experimental tasks.

Summary

The above information suggests that the three groups of children had approximately the same chronological age, the same number of boys and girls, similar psychological and physical health, and the same school grade level. They were significantly different in average intelligence. Eighty-five per cent of the fathers of high I.Q. children were in professional, or managerial and office occupations, whereas 80 per cent of the fathers of the low intelligence group were in unskilled, semi-skilled, skilled, and service occupations.

It is judged that these three groups were typical of children of their intelligence in communities of similar size, socio-economic level, and cultural conditions; and that the experimental conditions had similar motivational value for all groups.

The Population of Study Four

The sample of thirty children divided into three groups of ten by levels of intelligence and used in 1957 and 1959 laboratory studies of perception, while carefully drawn, was necessarily small in order to permit the individual laboratory testing of each child according to the various conditions of the designated tasks.

The findings of this intensive Madison Laboratory Study needed to be placed in a more adequate educational and population context and so it became necessary to design a comprehensive field study which would 1) permit more adequate sized groups of children to be identified at each intelligence level, 2) validate representativeness for a comprehensive residential and industrial urban area, and 3) balance the factors of enriched and impoverished educational background in respect to both schools and neighborhoods. It was argued that if the findings of the intensive study (Madison) would have some congruence with those of the field study (Milwaukee), then greater confidence can be had in the major findings of the total study.

In cooperation with members[8] of the central office staff of the Milwaukee Public Schools, eight school areas were identified which were judged representative of the different student populations, ethnic, cultural, economic conditions, and educational resources of the elementary schools of Milwaukee, Wisconsin.[9]

The following procedure was devised for determining the eight representative areas for Milwaukee. Five factors were discussed and agreed upon as being important in accounting for distinctive areas in the Milwaukee community. These factors were: 1) ethnic and racial composition of population (Negro, Porto Rican, German, Polish, and general); 2) industrial, business, or residential use of area; 3) homogeneity or heterogeneity of cultural and occupational backgrounds of people (community consciousness); 4) presence of unusual topographical, social, or legal barriers to social and economic movement; and 5) the location of pupil populations.

The Chairman of the research project then interviewed five members of the central staffs who had lived all their lives in Milwaukee, who had general supervisory responsibility for all of the schools in the community, who were familiar with the social and economic studies and conditions of the total community and who had first hand contact with the communities in which schools were placed. The interview procedure was to place the sheet with the five factors which might distinguish the geographic and cultural areas in Milwaukee before the interviewee with an outline map of Milwaukee with streets and salient geographic features noted. The instructions were to take a pencil and mark out as many distinctive areas as the interviewee could identify. If the time of the interview was not

---

[8]Dr. Dwight Teel, Assistant Superintendent in charge of Curriculum and Instruction; Miss Lillian Paukner, Director of Upper Grade Curriculum and Instruction; Miss Florence C. Kelly, Director, Primary Curriculum and Instruction; Dr. Eldon C. Bond, Division of Special Services (Assistant Superintendent); and the supervisory staff of the elementary schools.

[9]Milwaukee has 117 elementary schools serving a community of 367,392 people.

enough to satisfy the interviewee that he had properly identified these areas, he was allowed to take the materials overnight with the caution that he was not to discuss his outlining of areas with anyone.

The five maps obtained by this process were compared and all showed agreement on the general location of eight major distinctive areas in Milwaukee. Variation among maps were not in general location but in details of boundary lines, introduction of smaller areas with larger ones, and in dividing up large suburban residential areas. There was high agreement among the independent judgments of these five judges as to the representativeness of the identified eight regions.

The next step was to select one elementary school to represent each of these eight areas. The supervisory staff of the Milwaukee Public Schools, after considerable discussion and evaluation of tentative choices, agreed on eight schools in the separate areas as representing the full range of educational practice, pupil population, school physical plants, and the staff and leadership in the Milwaukee Elementary Schools.

These eight cultural, geographic, and economic areas, therefore, were considered representative of the range of conditions which differentiated the student populations and educational programs of elementary schools of the city. On the basis of these areas, eight elementary schools[10] were selected to constitute the total sample of study 4. The sample was assumed to be representative of the total school population of Milwaukee.

The bright population was obtained from the two special groups of children at the Maryland Avenue School (Miss Florence McCabe, Principal) and the William T. Sherman School (Miss Ruth E. Gruel, Principal), transported from the eight

---

[10]The eight elementary schools with their principals are: James Whitcomb Riley School, Mr. Raymond W. Ross, Principal; Forest Home Avenue School, Mr. Howard Williams, Principal; Henry L. Palmer School, Miss Margaret Borkowski, Principal; Mound Street School, Miss Ruth A. Foley, Principal; Twenty-Fourth Street School, Mr. George J. Kozina, Principal; Fifty-Third Street School, Mr. Paul H. Jacks, Principal; Cass Street School, Mr. James Burns, Principal; and Oklahoma Avenue School, Mr. Duane J Windau, Principal.

areas for a special program for the gifted. The sixty children included in the study were selected at random from this population and were assumed to be representative of all the gifted children of this age in Milwaukee.

The normal population of 60 children was selected by random numbers from all the 10 and 11 year olds of the eight selected schools who had complete intelligence test data and whose I.Q.'s were between 95 and 105.

The dull population of 60 children was selected by random numbers from the special classes for children with I.Q.'s of 80 and below in the selected schools. Modification was made within this group to match children of a given sex and age characteristic in order to reduce the travel between schools for the research assistants administering tests.

After the testing had been done it was discovered that 25 subjects in the dull group wrote manuscript and 35 wrote cursive. No manuscript scale had been devised to evaluate the legibility of the manuscript samples; therefore, the study had to be limited to 35 cursive writers. Five of these subjects in addition produced incomplete records so that the final sample used for legibility comparison was limited to 30 cursive writers for each of the three intelligence groups.

Random numbers were used to reduce the 50 bright and 60 normal groups to 35 each. Since the 60 dull subjects had been selected by random numbers from the dull children in eight selected schools, it was assumed that these 35 dull cursive writers were representative of all children in this intelligence classification in Milwaukee. There seems no reason to question this assumption. It is likely that the range of educational, social, and economic conditions which affect children of these intelligence levels are represented in this Milwaukee sample.

Intelligence Tests Used.

The WISC test was used on all children in the high and dull groups who also participated in the Klausmeier study. The California Test of Mental Maturity, Lorge-Thorndike, Kuhlman-Anderson group tests were used with the Binet, as an individual test, on all others. The average I.Q. in the three groups was 134 (high), 100 (average), and 68 (dull).

Test Situations

In the 1957 and 1959 Madison testing for the perception phase of the study, all subjects were tested in the University Laboratory to which each child had been transported. All testing was done by either a member of the research committee or a trained research assistant.

The Milwaukee population was tested in a room in each school set aside for the use of our research assistants. This room was private, included a table with two chairs, and individual children would come to this room as rapidly as the previous subject returned to his grade group. Because of the random selection of subjects, all eight schools participated.

All testing was done by research assistants transported to Milwaukee for this purpose.

Test Materials for Perception Study.

The Madison 1957 and 1959 testing included the collection of motor control, GSR, EMG, and point pressure handwriting records during the controlled writing phase. The equipment and experimental setting for the collection of these records is described in Chapter IV.

The complete directions and record forms for the collection of perception records in 1957 and 1959 are found in Appendix D (1957) and Appendix E (1959).

The directions and record forms for the Milwaukee Perception Study are found in Appendix F. Each research assistant had a pack which included the following material:

1. Sets of writing scales varying as to size and slant of letters.

2. A list of names of selected subjects.

3. A set of instructions and record forms.

4. A set of envelopes to place the completed cards and record for each subject.

5. A supply of 5x8 cards with "The quick brown fox jumps over the lazy dog." printed on each card and with lines. One card was used for each writing.

6. A stop watch for recording time for the motor coordination test and for checking to see if the subject wrote at least 10% faster than "usual" under the "fast" conditions.

7. A supply of 5x8 cards for the motor coordination test.

8. The experimental pen used for all subjects on the writing task.

9. Pencils for the coordination test.

10. Large envelopes to hold all materials.

The Handwriting Task Series

The handwriting tasks were designed to explore the kinds of handwriting models children set for themselves. Did they see their own handwriting in the same way as did adults when the handwriting was related to the same handwriting scale (normative model)? Did children of different intelligence distinguish between their own handwriting and how they would like their handwriting to look (aspiration model)?

In an effort to get at these questions, a series of tasks which were essentially the same for studies 1 through 4 were designed. The scale items and the subject's own samples all used the standard handwriting task: "The quick brown fox jumps over the lazy dog." Those who could only write manuscript used an abbreviated task, "The fox jumps."

Before each subject engaged in the experimental tasks, he was asked to write his name, address, and school. The slant and size characteristics of his

writing were used by the experimenter to select an appropriate set of scale items with writing characteristics similar to the subject's to be used in the testing. Thus, selected sets of scale items representing five levels of legibility were shuffled and placed before him in random order. The subject was allowed to determine his own scale ordering and to make judgments about a sample like his own and one he would want his own to look like without the bias of a determined adult ordering of this series of scale values. (Instructions to the experimenter are found in Appendix E.)

The nature, sequence, and repetition of certain tasks will be clarified by reference to Table 9.

Each subject engaged in the same sequence of six basic steps (A, B, C, D, E, and F), plus a seventh set (G) if his ordering of samples was not correct; i.e., did not agree with the adult ordering of scale values. The specific tasks within each step are numbered in special order, one through 11.

Examination of this table shows that each subject made certain judgments about his own perceived writing <u>before</u> actually writing (Step A), and then wrote samples of his own under the different conditions of writing which comprised Step B; and finally made a series of judgments <u>after</u> he had produced his own samples for use as possible referents (Steps C, D, E, F, and if necessary, Step G). The testing pattern required the subject to identify from scale items, writing samples which were like his <u>normative</u> and <u>aspired to</u> perceptions of his own writing without <u>any</u> samples of his own writing available for referent. The next major step was to have the <u>S</u> actually produce samples which represented the quality and style dimensions of his own writing. The third step was to use these actual samples to re-explore the judgments made in Step A under controlled conditions.

One should note, therefore, that each task required <u>before</u> writing (A1, 2, and 3) is repeated <u>after</u> writing (D5, 6, and F9). Furthermore, judgments about

TABLE 9

NATURE AND ORDER OF TASKS USED IN THE PERCEPTION OF THE HANDWRITING TASK

I Pre-Writing Phase	II Writing Phase	III Post-Writing Phase	
A. S judges scale samples before writing.	*B. S writes his own normal, best, and fast sample (1957+58); N_1, F,B,N_2,R (1959); N_1,F, B,N_2 (1960).	D. S judges scale samples after writing.	
1. S picks sample like his own from randomized samples.	C. S orders his own written samples.	5. S picks sample like his own from randomized samples.	F. E reshuffles scale items and has:
2. S picks sample aspired to from randomized samples.	4. S orders his own samples (easiest-to-read to hardest-to-read), selecting his best and poorest sample.	6. S picks aspired to sample from randomized samples.	9. S rearrange scale items placing easiest-to-read on top and hardest-to-read on bottom (2nd ordering of scale items).
		E. S judges own samples against 1st ordering of scale samples.	G. If 1st ordering was incorrect, E places items in correct order, and S judges his own sample against correct ordering of items.
		7. S picks scale sample like his best writing.	10. S picks scale sample like his best writing.
		8. S picks scale sample like his poorest writing.	11. S picks scale sample like his poorest writing.

*N,B,F (3 samples used in 1957 and 1958 studies.

N_1,F,B,N_2,R (5 samples used in 1959--study 3).

N_1,F,B,N_2 (4 samples used in 1960--study 4).

the scale values of the subject's best and poorest writing against his own ordering of scale samples (E7, 8), when incorrect, could be compared with similar judgments against correct ordering (as determined by adult judges of scale values, G10, 11). The repetition of tasks in this way provided an opportunity to check the consistency of agreement in judgments under these conditions. This represented one major type of analysis conducted.

The second major type of analysis had to do with relative judgments made within a given ordering. Thus, the relative position of samples judged like one's own (normative) and one "you would like your handwriting to look like"(aspired to) could be examined within the first ordering (A1, 2) before writing, and within the second ordering (D5, 6) after writing. Similarly, the relative position of the subject's own best and poorest samples could be determined in the first ordering (E7, 8) and in the correct ordering of the scale values (G10, 11) as well as the determination of how the best sample was ranked in comparison with the poorest sample (highest, the same, lowest) in steps E and G.

The third major type of analysis had to do with the assignment of legibility values to each writing sample produced by each S and ranked in order of best to worst. These legibility values permited further analysis of the normative perceptions of each individual as indicated by the correspondence of writing sample N_1 ("write as you usually do") with sample N_2 ("write as you usually do") produced after writing samples "as fast as you can" and "the best that you can." This behavioral demonstration of normative behavior (or lack of it) was very striking in the Ss in study 3 where complicated electrodes were attached to hands, arms, and the neck of each individual and extensive records were obtained (10-15 minutes of testing) between each of the "usual" writing samples. The legibility values also made possible the examination of the legibility variation each subject would permit in his writing under the different conditions of writing.

These three major type of analyses form the basis for the findings of this portion of the study.

Methodology and Results

Two important methodology problems needed to be solved before analyses could be made and results reported. These problems had to do with 1) the development of a scoring procedure for dealing with the rearrangement of scale and writing samples in the different choice situations, and 2) the assignment of legibility values to the different writing samples of the Ss in the study.

The Scoring of Rearrangement of Scaled Items

The S was asked to arrange a set of scaled writing samples in order from easiest to hardest to read (best to poorest) at two different times in the test series. This ordering by the subject may be compared with adult ordering and with each successive attempt. This is illustrated in Table 10.

TABLE 10

ARRANGEMENT OF SCALED WRITING SAMPLES IN ORDER FROM EASIEST TO HARDEST TO READ: SUBJECT 4

Value	1st	D^2	2nd	D^2
1	1	0	2	1
2	2	0	1	1
3	5	4	4	1
4	3	1	3	0
5	4	1	5	1
ΣD^2			6	4
Score			2	3

In this table, the value column represents the adult ordering of the scale items or key. The first and second columns represent the placement of these scale items by subject 4 in the two ordering situations. The deviation squared

of his placement with the key is indicated in the D^2 column. The sum of these D^2 is 6 for the first ordering and 4 for the second. The score (rho) for each ordering is 2 and 3 respectively. This is essentially the ranked order correlation of the subject's ordering with the key multiplied by the number of things ranked.[11]

---

[11] $p(\text{rho}) = \frac{1 - 6 \Sigma D^2}{N (N^2 - 1)}$

This procedure is developed by Stanley in Ross and Stanley, Measurement in Today's Schools (3rd edition), pp. 454-455.

The apparent difficulty of scoring rearrangement items has discouraged most testers from using them. Actually, when a suitable table is available, the task is not formidable. One way to score each item is to compute the rho (p) between the student's responses and the teacher's key and to multiply this rho by N, the number of things ranked. Thus, the testee who ranks on a certain six-option item agrees completely with the key secures a rho of +1 and a score of 1 x 6 = 6 for the item. If this rho on another such item is 0, he gets 0 x 6 = 6 points credit for that item. By being completely misinformed and securing a rho of -1, he could learn (-1) x 6 = -6 points. The rearrangement item is just about the only type that takes misinformation explicitly into account.

By using the formula, score (number of things ranked) x rho the various values of p are treated as if they constituted an equal-unit scale. Differences at the high end of the r scale represent greater changes in relationship than do differences of like magnitude between low r's. In an attempt to compensate for these inequalities, Stanley prepared the accompanying table by means of the formula S = (number of things ranked) x rho squared = Np^2.

TABLE 11

D^2 TABLE FOR SCORING REARRANGEMENT ITEMS:
SCORE (NUMBER OF THINGS RANKED) x $(\text{Rho})^2$*

	Number of Things Ranked (N)						
Score	2	3	4	5	6	7	Score
7						0- 2	7
6					0	4- 6	6
5				0	2- 4	8-10	5
4			0	2	6- 8	12-16	4
3		0	2	4	10-12	18-22	3
2	0			6- 8	14-16	24-30	2
1		2	4- 6	10-12	18-24	32-40	1
0		4	8-12	14-26	26-44	42-70	0
-1		6	14-16	28-30	46-52	72-80	-1
-2	2			32-34	54-56	82-88	-2
-3		8	18	36	58-60	90-94	-3
-4			20	38	62-64	96-100	-4
-5				40	66-68	102-104	-5
-6					70	106-108	-6
-7						110-112	-7

*Look up D^2 in Appropriate Column.

This scoring procedure is used in the analysis of all ordering data unless otherwise indicated.

Assignment of Legibility Values to Writing Samples

Two major assignments of legibility values were made 1) to the N,F,B,N,R samples of the 20 children of the high and average groups of the 1959 Madison study (100 samples) and 2) to the N,F,B,N samples of the 105 children in the high, average, and dull groups of the 1960 Milwaukee study (420 samples). The procedure was as follows:

The 667 scaled writing samples of the sixth grade handwriting master scale were used as the basis for assigning legibility values in both projects. This scale is described in Chapter II.

The 100 items of the Madison study and the 420 items of the Milwaukee project were measured for size and slant characteristics according to methods devised by the Committee as described in Chapter II. For each of the slant size combinations, a scale with legibility values from 1 to 5 was drawn from the items in the master scale. With these master scale items as a referent, 75 judges assigned legibility values to each of the items in the Madison Study and 80 judges to the items in the Milwaukee study. This rating for the Madison items was done during the summer of 1960 with graduate students recruited from the University of Wisconsin summer sessions classes. The rating of the Milwaukee items was done during the fall of 1960 with 80 graduate and upper class University students. The values were assigned to one decimal place.

The instructions to the raters were as follows:

1. We are interested in legibility, the ease with which you can read a given sentence.

2. You will be rating each of these handwriting samples on a scale ranging from 1 to 5. Number 1 represents poor handwriting and number 5 represents very good handwriting. You will notice samples on the scale that have already been assigned a number. Consider this number when you make your judgment. Some of the scales will have five samples to aid

you in your judgment; some will have fewer than five. Judge the handwriting in the light of these samples but remember the sample may run higher or lower in quality than the scale samples shown you.

3. Your task is to give each of these samples of handwriting a number on the legibility scale 1 through 5. This number should be written to the first decimal place, i.e., 2.5, 3.0, 1.8. Consider the first sample. Decide the number that most approximates that sample and list it on your data sheet under the appropriate "blue tag" number. Continue with the rest of the samples in this order.

4. Do them as rapidly as you can, but be as accurate as possible.

The mean of the 80 ratings was taken as the legibility value of the experimental sample. These legibility values or scores have the advantage of being on a cardinal scale and thus can be compared directly and relatively.

Results of the Perception Studies

The results of the perception studies are organized by the major questions asked with the findings of the Milwaukee study presented first. The data from the other studies will be related to these findings as a basis for verifying the representativeness of these results.

Ordering of Handwriting Scale Samples

The Ss of all three groups in the Milwaukee study were asked to arrange scaled samples in order from easiest-to-read to hardest-to-read both before and after writing their own experimental sample. These two orderings were scored according to the procedure described previously. An analysis of variance design was used to compare the arrangement scores of the 35 subjects in each group. Since the sexes were not equally represented in the slow group, an analysis of sex differences was not attempted. The results of this analysis is shown in Table 12.

This analysis shows that the three ability groups differed significantly in correctly ranking the scale samples as shown in the table. Scores used in

TABLE 12

ANALYSIS OF VARIANCE AMONG BRIGHT, AVERAGE, AND DULL GROUPS IN ORDERING HANDWRITING SAMPLES CORRECTLY (Milwaukee)

		I	II	Total
IQ	H	117	125	242
	A	100	113	213
	D	67	67	134
Total		284	305	589

Summary

Source	df	Sum Sq.	Mean Sq.	F
IQ	2	89.27	44.635	10.991**
Sub/Gps	102	414.23	4.061	
Tasks	1	2.10	2.100	1.118
Tasks x IQ	2	1.23	.615	
Ss x Task/Gp	102	182.18	1.786	

**Significant at the .01 level.

the analysis of variance were based upon the sum of the total deviations, squared, for the first and second orderings: $\Sigma(D_{1,2})^2$. The observed differences appear to be related to intelligence, the bright group showing the highest agreement with the correct scale ordering, the dull group showing the least. Both bright and average groups showed some improvement in their second ordering, whereas the dull group showed none, as indicated in Columns I and II in Table 12.

The significance of the difference in performance was also examined by applying "t" tests. The results are shown in Table 13. If the .01 per cent level

TABLE 13

"t" TEST RESULTS FOR THE SIGNIFICANCE OF DIFFERENCE IN MEANS BETWEEN BRIGHT, AVERAGE, AND DULL GROUPS IN ORDERING HANDWRITING SAMPLES CORRECTLY

High-Dull
$t = 4.312$; $df = 68$; $P = < .01$; significant

High-Average
$t = 1.369$; $df = 68$; $P = > .05$; not significant

Average-Dull
$t = 3.149$; $df = 68$; $P = < .01$; significant

of significance is used as the criterion, it will be seen that the high-dull and average-dull groups differed significantly in their performance on this task.

The high-average group did not. The "t" test results thus support the findings of the analysis of variance. Thus, the original hypothesis that the three ability groups would perform equally well was not supported in this specific task.

The lack of difference in the performance of the high and average groups in ordering handwriting samples correctly in the Milwaukee study was also substantiated in the Madison 1959 data. In the high-average group comparison on the same question, $t = 1.7422$; $df = 18$; $P = <.10$; not significant.

To What Extent Do Bright, Average, and Dull Groups Agree With the Correct (Adult) Ordering of Handwriting Samples?

If complete agreement is the criterion, this question may be answered very simply by saying "very little." Specifically, the number and per cent of perfect agreements may be seen in Tables 14 and 15 for Milwaukee and Madison data.

TABLE 14

NUMBER AND PER CENT OF COMPLETE AGREEMENT IN ORDERING BY BRIGHT, AVERAGE, AND DULL GROUPS (Milwaukee)

Group	N	Correct Agreement (First Ordering)		Correct Agreement (Second Ordering)	
		N	%	N	%
Bright	35	9	26	12	34
Average	35	6	17	6	17
Dull	35	4	11	3	9

TABLE 15

NUMBER AND PER CENT OF COMPLETE AGREEMENT IN ORDERING BY BRIGHT AND AVERAGE GROUPS (Madison)

Group	N	Correct Agreement (First Ordering)		Correct Agreement (Second Ordering)	
		N	%	N	%
Bright	10	2	20	2	20
Average	10	4	40	3	30

However, these figures are substantially increased if the interchange of two adjoining scale items (scored as a "4" instead of a perfect "5") is allowed and combined with perfect scores. In this case, the figures would be those indicated in Tables 16 and 17.

TABLE 16

NUMBER AND PER CENT OF NEAR AGREEMENT IN ORDERING BY BRIGHT, AVERAGE, AND DULL GROUPS (Milwaukee)

Group	N	Correct Agreement (First Ordering)		Correct Agreement (Second Ordering)	
		N	%	N	%
Bright	35	20	57	22	63
Average	35	16	49	19	54
Dull	35	11	31	8	23

TABLE 17

NUMBER AND PER CENT OF NEAR AGREEMENT IN ORDERING BY BRIGHT AND AVERAGE GROUPS (Madison)

Group	N	Correct Agreement (First Ordering)		Correct Agreement (Second Ordering)	
		N	%	N	%
Bright	10	5	50	9	90
Average	10	6	60	6	60

Note, however, that on either basis of scoring, the relative positions of the bright, average, and dull groups remain the same. On the more severe criterion of perfect agreement, only the bright group showed such agreement in from one-fourth to one-third of the judgments made. The Madison high-average groups, however, were much alike with the average somewhat better. On the more lenient basis of scoring, however, the percentages clustered in the mid-50 per cent range for the bright and average groups and in only one case (bright, second ordering) approached agreement in 63 per cent (Milwaukee) and 90 per cent (Madison) of the judgments made.

It would thus appear, in answer to the two basic questions raised, that these groups of children ordered handwriting samples differently, and that their judgments of correct ordering were on the whole different from those of adults.

These results present some interesting problems regarding the ability of middle grade elementary school children to perceive a continuum of quality in handwriting even when this continuum is limited to five levels of difference. (Adult criterion was 100% agreement among 10 judges as to placement of scale samples; there was, however, no control in this scale as to the amount and equality of the quality difference between the five scale samples. Inspection, however, shows that misplacement by children occurred at all levels in the scale items.)

The formation of an internalized generalized model of quality by a child which he can use to arrange samples of handwriting in a given quality order obviously does not develop rapidly (children here were in the fifth and sixth grades and had had five or six years of writing instruction). It seems clear that broad level or range of intelligence is a condition of importance here. In all cases, the bright and average groups were much more successful in this task than were the dull group. Yet the difference between the average and the bright group was not significant in either the Milwaukee or Madison studies. Only when the "near agreement" criterion was used did the bright group become distinctly superior in performance on this task.

Too much should not be made of these results, and subsequent performance on related perception tasks should indicate the confidence to be placed in this specific finding. It does suggest, however, an area of development in skill learning not often recognized, that of a generalized quality perception within which a child's "normative" and "aspired to" models will have a recognized place. Efforts to develop this quality perception in the skill learning behavior may be equal to if not more important than the development of skill in

specific letter and word formation. Certainly what has been called "evaluation" as an important part in learning will be based in a large part by an individual on his generalized quality perception of the nature of the learning in which he is engaged. The giving of an important role to the development of this kind of perception in skill learning would change significantly present educational practices in this area of school instruction.

Judgments on Specific Tasks
Relative to the Subject's Ordering
and the Correct Ordering of Samples.

The repetition of certain tasks noted on Table 9 made it possible to examine the agreement in performance on successive comparable tasks. A chi-square analysis was used to study the relative performance of the bright, average, and dull groups in nine aspects of the handwriting tasks. These nine aspects of the four major tasks are:

A. Does each S place a sample "like mine" in the same way in an array of scaled items (1) before and (2) after writing his experimental samples?

B. Does each S place an aspired to sample in the same way in an array of scaled items (3) before and (4) after writing?

C. Do children place their "like mine" and aspired to samples in the (5) same relative position in their ordering? (The relationships can be higher, lower, and equal to.)

D. Do children place their 'best' and their 'poorest' samples in the same relative position in their (6) own ordering and in the (7) correct ordering of samples?

E. Do children place their 'best' and poorest samples in the same position relative to each other (8) in their own ordering and (9) in the correct ordering of scale samples?

It should be noted, however, in connection with certain aspects of the following analysis that the dull group was confronted with especially difficult decisions because only 37 per cent of this group produced three completely accurate reproductions of the entire task sample, whereas all others of this group succeeded in making only incomplete or partial reproductions. Since it was

found in our previous work that comparisons of part with whole handwriting samples are difficult, this group was undoubtedly greatly handicapped in making certain required judgments.

Do children place the sample "like mine" in the same relative position in their first ordering (1) and in their second ordering (2) of samples before and after writing? The results of the chi-square analysis shown in Table 18 indi-

TABLE 18

AGREEMENT ON SUCCESSIVE ORDERINGS OF "LIKE MINE" SAMPLES
(Milwaukee)

	B	A	D	Total
1 > 2	8 .16	7 .12	7 .20	22
1 = 2	33 .66	46 .76	15 .43	94
2 > 1	9 .18	7 .12	13 .37	29
Total	50	60	35	145

$X^2 = 11.976$
$P < .02$

cate that the agreement in two orderings is significantly different from that accounted for by chance. It may be observed that the bright group identified its "like" sample in the same relative position in the two orderings in much higher a per cent of their judgments than did the average and dull groups. The latter groups appeared to be as prone to place "like" samples in a higher position in the second ordering as to place it in the same relative position. Thus, only the bright children, who were shown to be most accurate and stable in ordering samples in the preceding section, appeared to be very consistent in selecting samples "like mine" in two successive orderings of handwriting scale samples.

This may also suggest that the bright group had a clearer percept of their own handwriting initially (before writing) which was not appreciably altered by the task of writing their own samples before making a judgment of the sample "like mine" in the second ordering.

The examination of the Madison data shows (see Table 19) that the bright

TABLE 19

AGREEMENT ON SUCCESSIVE ORDERINGS OF "LIKE MINE" SAMPLES
(Madison)

	B	A	Total
1 > 2	3 .3	2 .2	5
1 = 2	4 .4	3 .3	7
2 > 1	3 .3	5 .5	8
Total	10	10	20

$X^2 = 2.000$
$P < .70$

and average groups did not differ from chance in their successive ordering of their "like mine" samples. The chi-square indicates that the chance hypothesis cannot be discarded. Yet examination of the individual cases observed that if the "near" criteria of one level placement difference was used all of the bright children would have met this criterion. Only seven of the average group would have met it.

Do children place the aspired to sample in the same relative position in their first ordering (3) and their second ordering (4) of samples before and after writing? The chi-square analysis in Table 20 shows that agreement on successive orderings of aspired to samples was not a chance agreement. Both high and average groups showed a comparable tendency to select the same relative position for aspired to scale samples in their successive orderings, whereas the dull group approximated a chance distribution in their judgments. Each group, however, made a higher percentage of choices in agreement (that is, 1 = 2) with the same relative position of the aspired to scale sample than in non-agreement. These findings suggest greater stability among children in selecting the same relative position of the aspired to sample in successive orderings of scale samples than in the selection of their "like mine" sample. This suggests that the

TABLE 20

AGREEMENT ON SUCCESSIVE ORDERINGS OF "ASPIRED TO" SAMPLES
(Milwaukee)

	B	A	D	Total
I > 2	8 .16	7 .12	7 .20	22
1 = 2	33 .66	46 .76	15 .43	94
2 > 1	9 .18	7 .12	13 .37	29
Total	50	60	35	145

$X^2 = 11.976$
$P < .02$

normative perception of one's own handwriting does not stabilize for individual children as quickly as does their picture of what they would like their handwriting to look like.

Table 21 shows the findings for the bright and average groups on the same

TABLE 21

AGREEMENT ON SUCCESSIVE ORDERING OF
"ASPIRED TO" SAMPLES
(Madison)

	B	A	Total
1 > 2	2 .2	1 .1	3
1 = 2	6 .6	8 .8	14
2 > 1	2 .2	1 .1	3
Total	10	10	10

$X^2 = 13.667$
$P < .01$

question in the Madison data. If any, there is even a greater tendency for the Madison children to have chosen the same relative position for the aspired to scale in their successive orderings.

Do children place their "like mine" and aspired to sample in the (5) same relative position in their ordering of samples? Since the children made the two judgments before writing and before the first ordering of the scale samples and again after writing and making a second ordering of scale samples, it is possible to examine the relative position within each ordering of their "like mine" and aspired to samples. Chi-square analysis indicated in Tables 22 and 23 for

TABLE 22

RELATIVE POSITION OF "LIKE MINE" AND ASPIRED TO SAMPLES IN THE FIRST ORDERING OF SCALE SAMPLES (Milwaukee)

	B	A	D	Total
M > A	11 .22	2 .03	7 .20	20
M = A	19 .38 (34*)	24 .40 (30*)	8 .23 (14*)	51
A > M	20 .40	34 .57	20 .57	74
Total	50	60	35	145

*Per cent of "like mine" and aspired to samples placed in the highest rank order (rank 1).

$X^2 = 11.597$

$P < .05$

TABLE 23

RELATIVE POSITION OF "LIKE MINE" AND ASPIRED TO SAMPLES IN THE SECOND ORDERING OF SCALE SAMPLES (Milwaukee)

	B	A	D	Total
M > A	8 .16	10 .17	8 .23	23
M = A	24 .48 (40*)	26 .43 (38*)	6 .17 (14*)	56
A > M	18 .36	24 .40	21 .60	63
Total	50	60	35	145

*Per cent of "like mine" and aspired to samples placed in the highest rank order (rank 1).

$X^2 = 10.428$

$P < .05$

the first and second orderings, respectively, that the agreements in relative position differ significantly from that expected by chance, though to a lesser degree than that found for agreement in the relative position of these judgments in successive orderings. A first glance at these tables would suggest that average and dull children had a higher level of aspiration in this task than did bright children. This, however, may be an artifact of the task. (There was no higher scale value left to choose.) This was not true to as nearly as great an extent for the dull group which seldom made both top-ranked samples their choices. Thus, the dull group could and did tend to show a higher aspiration level. A substantial per cent of the bright and average groups could not make a higher judgment than M = A since they had reached the top of the scale. It is perhaps best simply to consider the combined percentages for M = A and A > M in noting that these children in both orderings showed considerable consistency in placing their aspired to sample at either the same or at a higher position than the sample judged "like mine." In this sense, they appeared to interpret the direction "show me the one you would like your handwriting to be like" in accordance with the intent of the question.

Tables 24 and 25 present the findings of the Madison study. The first ordering supports the Milwaukee data. The second ordering, however, would suggest that chance is a better explanation than the previous finding. The average group is the only one which corresponds to the previously determined pattern (A > M).

<u>Do children place their best writing samples and their poorest writing samples respectively in the same relative position in their own ordering (6) and the correct ordering (7) of samples</u>? When children were asked to perform this task first with respect to their own ordering of handwriting samples and then against the correct ordering, the agreement in these parallel tasks approximated chance agreement, as the chi-square analysis in Tables 26 and 27 show.

TABLE 24

RELATIVE POSITION OF "LIKE MINE" AND ASPIRED TO SAMPLES
IN THE FIRST ORDERING OF SCALE SAMPLES
(Madison)

	B	A	D	Total
M > A	4 .40	1 .10	1 .10	6
M = A	4 .40 (35*)	4 .40 (51*)	1 .10 (50*)	9
A > M	2 .20	5 .50	9 .90	16
Total	10	10	10	30

*Per cent of "like mine" and aspired to samples placed in the highest rank order (rank 1).

$X^2 = 18.667$

$P < .01$

TABLE 25

RELATIVE POSITION OF "LIKE MINE" AND ASPIRED TO SAMPLES
IN THE SECOND ORDERING OF SCALE SAMPLES
(Madison)

	B	A	D	Total
M > A	4 .40	1 .10	2 .28	7
M = A	3 .30 (50*)	4 .40 (60*)	2 .28 (52*)	8
A > M	3 .30	5 .50	3 .44	12
Total	10	10	7	27

*Per cent of "like mine" and aspired to samples placed in the highest rank order (rank 1).

$X^2 = 3.8333$

$P < .30$

TABLE 26

RELATIVE POSITION IN FIRST AND SECOND
ORDERINGS OF BEST SAMPLE
(Milwaukee)

	B	A	D	Total
1 > 2V*	9 .26	19 .40	11 .35	39
1 = 2V	15 .43	8 .17	7 .23	30
2V > 1	11 .31	20 .43	13 .42	44
Total	35	47	31	113

*V = Correct value scale.

$X^2 = 8.309$

$P < .10$

TABLE 27

RELATIVE POSITION IN FIRST AND SECOND
ORDERINGS OF POOREST SAMPLE
(Milwaukee)

	B	A	D	Total
1 > 2V*	10 .29	13 .28	8 .26	31
1 = 2V	10 .29	19 .40	13 .42	42
2V > 1	15 .42	15 .32	10 .32	40
Total	35	47	31	113

$X^2 = 1.799$

$P < .80$

It may be inferred that this task which involved actually studying their own sample and deciding where it belonged on a handwriting scale made up of samples written by others was sufficiently difficult, if not confusing, to produce inconsistencies in the judgments of the majority of children.

The Madison data is incomplete on this question but an analysis of the completed records show that the Milwaukee findings are supported.

Do children place their best and poorest samples in the same relative position to each other in their own ordering (8) and in the correct ordering (9) of scale samples? The chi-square analysis of this question is shown in Tables 28 and 29. Both relative positions assigned in their own, the first ordering, and those judgments made with respect to the correct ordering differed significantly from a chance distribution. The trend for children in all groups in both orderings was clearly in the direction of ranking their best higher than their poorest. It should be noted, though, that the dull group in both orderings made a higher percentage of P > B than did bright and average groups. Children in the latter groups frequently commented that they felt their best and poorest samples were sufficiently alike to make it difficult to say that one was more readable than the other.

TABLE 28

RELATIVE POSITION OF BEST AND POOREST SAMPLES IN FIRST ORDERING (Milwaukee)

	B	A	D	Total
P > B	8 .16	13 .22	10 .29	31
P = B	20 .40 (28*)	12 .20 (07*)	8 .22 (02*)	40
B > P	22 .44	35 .58	17 .49	74
Total	50	60	35	145

*Per cent of best and poorest samples placed in the highest rank order (rank 1)

$X^2 = 27.2500$

$P < .01$

TABLE 29

RELATIVE POSITION OF BEST AND POOREST SAMPLES IN CORRECT ORDERING (Milwaukee)

	B	A	D	Total
P > B	5 .14	7 .15	12 .39	24
P = B	11 .31 (17*)	11 .23 (06*)	4 .13 (00*)	26
B > P	19 .55	29 .62	15 .48	63
Total	35	47	31	113

$X^2 = 9.514$

$P < .05$

The figures shown in parenthesis for P = B show that a much more substantial percentage of the bright group placed both their best and poorest samples in the highest rank order than did the average and dull groups. In this case, such children could not rate B > P, a scale artifact which theoretically lowered the possible per cent in the latter category. If the percentages for P = B and B > P are combined, a considerably higher per cent of bright and average groups judged their best sample at least equal to or better than their poorest than did the dull group.

Tables 30 and 31 for the Madison study support the findings for the Mil-

TABLE 30

RELATIVE POSITION OF BEST AND POOREST SAMPLES IN FIRST ORDERING (Madison)

	B	A	Total
P > B	3	1	4
P = B	4	0	5
B > P	1	5	6

$X^2 = 8.3333$
$P < .05$

TABLE 31

RELATIVE POSITION OF BEST AND POOREST SAMPLES IN CORRECT ORDERING (Madison)

	B	A	Total
P > B	0	2	2
P = B	8	2	10
B > P	2	6	8

$X^2 = 15.3333$
$P < .01$

*Per cent of best and poorest samples placed in the highest rank order (rank 1).

waukee data. More than chance operated in the choices and the tendency was to place the best sample ahead of the poor. This tendency is most clear in the average group. The bright group data is obscured because of the many placements of the best and poorest samples in the highest rank (1) by the bright group.

The Relation of Perception and Legibility

The development of the master set of 677 scaled handwriting items measured for size and slant provided the authors with a tool for actually examining the differences which existed in legibility between the handwriting samples being judged by the three intelligence groups. (Description of the scaled items and

the procedure for assigning legibility values to the experimental samples has been discussed in the chapter on legibility and in the first portion of this chapter.)

General Relation of Legibility to Conditions of Experiment

Before examining the relationship between perception and legibility, it was necessary to examine the general relationship of the quality of legibility to I.Q., sex, and the conditions (N,F,B,N) in both the Milwaukee and the Madison samples. The analysis of variance was done for both samples and is presented in Tables 32 and 33.

TABLE 32

ANALYSIS OF VARIANCE: RELATIONSHIP OF LEGIBILITY TO I.Q., SEX, AND CONDITIONS (Milwaukee)

Source	df	SS	MS	F
IQ	2	4491.25	2245.63	26.96**
Sex	1	694.44	694.44	8.34**
IQ x Sex	2	38.32	19.16	
Ss/Groups	84	6995.31	83.28	7.92**
Conditions	3	256.89	85.63	8.15**
Cond x IQ	6	51.70	8.62	
Cond x Sex	3	13.22	4.41	
Cond x Sex x IQ	6	173.83	28.97	2.76*
Cond x Ss x Groups	252	2649.10	10.51	
	359			

** P <.01
* P < .05

TABLE 33

ANALYSIS OF VARIANCE: RELATIONSHIP OF LEGIBILITY TO I.Q., SEX, AND CONDITIONS (Madison)

Source	df	SS	MS	F
IQ	1	.5983	.5983	
Sex	1	5.1309	5.1309	8.432*
IQ x Sex	1	.5883	.5883	
Ss/Groups	16	9.7362	.0685	18.6656**
Conditions	3	.8983	.2994	9.184**
Cond x IQ	3	.0077	.0026	
Cond x Sex	3	.1545	.0515	1.580
Cond x Sex x IQ	3	.2294	.0765	2.346
Cond x Ss x Groups	48	1.5633	.0326	
	79			

** P <.01
* P <.02

The results of this analysis showed that for the Milwaukee sample there were significant relationships between legibility and I.Q., sex, and conditions at the .01 level of confidence; and conditions, sex, and I.Q. at the .05 level. For the Madison sample, a significant relationship was found for sex and conditions. This general finding is important in providing a background for dealing with the individual questions which follow.

Four major questions were asked as a basis for dealing with these data:

1. How accurately do bright, average, and low I.Q. children order their handwriting samples in relation to legibility scale judgments by adults?

2. What is the difference in legibility of the two normative samples as a consequence of I.Q.?

3. What is the range of quality in handwriting in the three conditions--best, normal, and fast--as a consequence of I.Q.?

4. What is the relationship between perception, legibility, and I.Q.?

<u>How accurately do bright, average, and low I.Q. children order their handwriting samples in relation to legibility scale judgments by adults?</u>

Each child was asked to take his normal, fast, and best handwriting samples produced under the experimental conditions and order them in relation to their legibility. This ordering was compared to the actual legibility values assigned to these samples by procedures described in the first part of this chapter.

Since we wished to examine this ordering in relation to sex (our girl writers had produced consistently better writing in previous investigations), the population of each group compared was composed of 30 individuals each, 15 boys and 15 girls, selected randomly within the definitions predescribed by the design of the study. Table 34 shows the result of this examination.

The reader will remember that lack of correspondence between ordering of children and adult scaling is indicated by a low score, high correspondence approaches a score of 3. These scores are summed and averaged for each intelligence group and by sex.

TABLE 34

CORRESPONDENCE OF THE ORDERING
OF BRIGHT, AVERAGE, AND DULL CHILDREN OF
THEIR HANDWRITING SAMPLES FOR LEGIBILITY AND ADULT SCALING
(Milwaukee)

		IQ				
		H	A	D	Total ΣD^2 Score	Average
Sex	F	19	25	1	45	1.00
	M	9	17	3	29	.64
Total ΣD^2		28	42	4	74	.82
Average		.94	1.40	.14		
N		30	30	30		

Scores are in terms of D^2 units.
Optimum score 3 (highest correspondence).
N = 90.

Again one is struck with the internal consistency of the data when compared with previous results. The high and the average groups are significantly more accurate than the low, the average again is superior to the other groups, and the girls are superior to the boys in all groups excepting the low where some slight advantage is with the boys.

Table 35 presents the Madison data on the same question. This table supports

TABLE 35

CORRESPONDENCE OF THE ORDERING
OF BRIGHT AND AVERAGE CHILDREN OF THEIR
HANDWRITING SAMPLES FOR LEGIBILITY WITH ADULT SCALING
(Madison)

		IQ			
		H	A	Total ΣD^2 Score	Average
Sex	F	5	7	12	1.20
	M	7	7	14	1.40
Total ΣD^2		12	14	26	
Average		1.20	1.40		
N		10	10		

Scores are in terms of D^2 units.
Optimum score 3 (highest correspondence).
N = 20.

the equal and better accuracy of the average group over the bright group but does not support the superiority of the girls over the boys (boys were equal or better).

The two exploratory studies in 1957 and 1958 with Madison and with Laboratory children showed that normal and bright children could place their normal and best sample on a scale in greater agreement with adult judges than could the dull children. Thus, all the evidence on this point is consistent. The increased power of our evaluation instruments merely reinforce what the first exploratory studies suggested.

<u>What is the difference in the quality of the two normal conditions (usual writing) as a consequence of I.Q.?</u>

The subjects in the 1959 Madison and Milwaukee studies wrote the sequence of normal^1, fast, best, and normal^2 samples under the indicated conditions. These data provided an opportunity to examine the normative perceptions of the three groups differentiated by intelligence. The capacity of an individual child to "write as you usually do" (N^1), and then write "as fast as you can" (F), "as well as you can" (B), and then another "as you usually do" (N^2) should indicate the extent to which this normative perception of his own writing is stabilized for that individual. Table 36 presents the data on this question for the Milwaukee sample and Table 37 for the Madison sample.

The data in these tables are in terms of the average ranges of legibility values between the N^1 and N^2 samples for each group and each sex. In this case, a low score means high normative perception and a high score means low normative perception.

The examination of the Milwaukee data reveals little difference in the relative performance of high, average, and low intelligence groups and by sex on this task. There is some slight suggestion that the average group and the boys were more consistent in their normative perceptions than were the other groups and girls.

TABLE 36

THE AVERAGE DIFFERENCE IN QUALITY BETWEEN TWO USUAL WRITING SAMPLES AS A CONSEQUENCE OF IQ AND SEX (Milwaukee)

		IQ			
		H	A	D	Total
Sex	F	.26	.31	.34	.30
	M	.32	.24	.26	.27
Total*		.29	.28	.30	.29
N		30	30	30	90

*In terms of range averages.

TABLE 37

THE AVERAGE DIFFERENCE IN QUALITY BETWEEN TWO USUAL WRITING SAMPLES AS A CONSEQUENCE OF IQ AND SEX (Madison)

		IQ		
		H	A	Total
Sex	F	.17	.17	.22
	M	.09	.30	.20
Total*		.13	.29	.21
N		10	10	20

*In terms of range averages.

The Madison data suggest a clear-cut difference between the bright and average groups--the bright group (and boys in the bright group) apparently have stabilized this kind of perception of their writing. The total difference in sexes is not significant other than to point out that the relative good performance of the bright boys is balanced by the relatively bad performance of the average boys. This result is interesting in that the Madison experimental conditions involved elaborate electrode connections to each S to secure electromyographic records.

What is the range of quality in handwriting in the three conditions--best, normal, and fast--as a consequence of I.Q.?

Tables 38 and 39 present the Milwaukee and Madison data on this question. The data reveal that there is no difference in the range of quality of handwriting in the three conditions of best, normal, and fast as a consequence of I.Q.

TABLE 38

RANGE OF QUALITY IN THE THREE CONDITIONS OF WRITING AS A CONSEQUENCE OF IQ AND SEX (Milwaukee)

		IQ			
		B	A	D	Total
Sex	F	.55	.65	.43	.54
	M	.59	.57	.53	.56
Total*		.57	.61	.48	.55
N		30	30	30	90

*In terms of range averages.

TABLE 39

RANGE OF QUALITY IN THE THREE CONDITIONS OF WRITING AS A CONSEQUENCE OF IQ AND SEX (Madison)

		IQ		
		H	A	Total
Sex	F	.34	.38	.36
	M	.37	.38	.38
Total*		.36	.38	.37
N		10	10	20

*In terms of range averages.

in the bright and average groups and by sex in both the Milwaukee and Madison studies. There is, however, a significant difference between the performance of these two groups and the low group in intelligence in the sense that the range of their writing performance was less. In this low group also, the performance of the girls was even more restricted than the boys.

These findings are further supported by a study of the range of legibility in the samples of each intelligence and sex group in the Milwaukee data, as shown in Table 40. This table indicates what one might expect. The high intelligence group had the highest quality sample, wrote in the higher and broader range of quality; the lower group wrote over the lower and narrower range of quality and had the poorest sample; while girls had the best and the poorest sample and thus wrote over a wider range of quality than did the boys. These findings are somewhat of a surprise--especially in respect to range--the girls being greater than the boys.

TABLE 40

RANGE OF LEGIBILITY OF THE WRITING SAMPLES
BEST, NORMAL, AND FAST BY INTELLIGENCE AND SEX*
(Milwaukee)

		IQ			
		H	A	D	Total
Sex	F	4.51-2.24	4.07-1.53	3.11-1.00	4.51-1.00
	M	3.71-1.54	3.55-1.44	3.13-1.29	3.71-1.29
Total		4.51-1.54	4.07-1.44	3.13-1.00	4.51-1.00

*Five point scale.

What is the relationship between perception, legibility and I.Q.?

Through procedures described in the first of the chapter, it is possible to obtain a perception index and a legibility index which can be used to dichotomize the 90 and 20 individuals of the Milwaukee and Madison studies into high perception and low perception groups.[12] The legibility index was used to separate the same groups into high and low legibility sections.[13] The two samples were already divided into three intelligence groups: high, average, and low.[14]

Table 41 presents three comparisons for the Milwaukee data. For each intelligence group it is possible to identify individuals with high perception and

---

[12]A perception index was computed by measuring the degree of agreement between the student's ordering of three of his handwriting samples and the legibility ratings assigned by adult judges of these three handwriting samples. In the ordering of these three handwriting samples, those individuals categorized as high perception showed perfect agreement with the adult judges, or showed one deviation in ordering from the adult judges' rating. Those individuals categorized as low perception showed two deviations in ordering or three deviations in ordering from the adult judges' ratings.

[13]A legibility index was derived by scoring the best sample of each subject's handwriting on the basis of a scale value from 1-5 (poorest to best) from the Wisconsin Legibility Scale. Those individuals categorized as high legibility had scores on the legibility scale of 3-5. Those individuals categorized as low legibility had scores on the legibility scale of 1-2.9.

[14]Those individuals categorized as high I.Q. had mental maturity scores ranging from 116-171. Those individuals categorized as average I.Q. had mental maturity scores ranging from 90-110. Those individuals categorized as low I.Q. had a range in mental maturity scores of 51-83. The population in each of the three I.Q. groups was 30.

TABLE 41

RELATIONSHIPS BETWEEN LEGIBILITY, PERCEPTION, AND INTELLIGENCE
(Milwaukee)

Perception	Legibility	IQ H	IQ A	IQ D	Total
High	High	.43 13	.43 13	.07 2	28
	Low	.30 9	.40 12	.43 13	34
Low	High	.10 3	.07 2	.07 2	2
	Low	.17 5	.10 3	.43 13	22
Total		30	30	30	

df = 6
$X^2 = 17.70 < .01$

with high and low legibility; also with low perception and with high and low legibility. These same decisions are made for the average and low intelligence groups.

The chi-square obtained shows that these relationships cannot be explained by chance alone. Some interesting comparisons can be made:

High I.Q.: 73% have high perception; 53% have high legibility.

Average I.Q.: 83% have high perception; 50% have high legibility.

Low I.Q.: 50% have high perception; 14% have high legibility.

Again the close correspondence that exists between the high and average intelligence groups both on perception and legibility is revealed. In all groups more individuals have higher perception scores than legibility scores. This is emphasized further when low legibility is used as the referent for the comparison.

High I.Q.: 47% have low legibility; 65% of these have high perception.

Average I.Q.: 50% have low legibility; 60% of these have high perception.

Low I.Q.: 83% have low legibility; 50% of these have high perception.

While this shows that the low I.Q. group had many more individuals with low legibility, about the same fraction of these individuals had high perception scores as had the bright and average groups.

Or, putting these findings in another way, if a child has high perception, he has a 50/50 chance of having high legibility. If he has low perception, there is 1/3 chance of having high legibility.

The Madison data, shown in Table 42, does not reveal any conclusive

TABLE 42

RELATIONSHIPS BETWEEN LEGIBILITY, PERCEPTION, AND INTELLIGENCE
(Madison)

Perception	Legibility	IQ H	IQ A	Total
High	High	.10 1	.40 4	5
	Low	.60 6	.50 5	11
Low	High	.20 2	0	2
	Low	.10 1	.10 1	2
Total		10	10	20

$X^2 = 2.90$ Not significant

findings.

While an explanation of the Madison data cannot discard chance alone as an explanation of these relationships, the data do suggest some interesting possibilities. Perhaps the most important part of this picture, however, is that 60% of the high intelligence group and 50% of the average group had a high perception-low legibility pattern. Or putting it another way, 86% of the high intelligence group with high perception had low legibility. There was no indication of this pattern (50/50), however, for the average group.

Summary and Discussion of Findings

The general findings of this chapter can be organized into three major headings:

I. General perception of differences in quality in handwriting by three groups differing in intelligence.

A. The three intelligence groups differed in their performance in ordering handwriting samples as to quality; the performance of the bright and average groups were not significantly different in respect to each other; the bright and average groups differed significantly from the performance of the dull group.

B. The bright and average groups ordered the scale samples more in accordance with adult placement than did the dull group. The per cent of correct placement for the bright and average groups ranged from 63 to 90%. The dull averaged around 25%.

C. The bright and average groups became more accurate on the second ordering; the dull group did not.

II. The consistency of two successive orderings of "like mine" (normative) and desired (aspired to); best and poorest samples by three groups differing in intelligence.

A. All intelligence groups differed in their two placements of their "like mine" samples and all were more consistent than chance. The bright group seemed to have a more stable perception of a sample like their own.

B. All intelligence groups placed their aspired to sample in two orderings more consistently than chance. The bright and average groups evidenced a more stable perception of what they would like their handwriting to look like than did the dull group. The aspired to perception seemed more stable than the perception "like mine."

C. All groups consistently placed their aspired to sample equal to or higher in quality than the sample "like mine."

D. All groups did not place their best and poorest writing samples on a scale in two orderings at a level better than chance.

E. All groups differed significantly from chance, however, in the relative placement of the best and poorest samples in two orderings. The trend in all groups was to rank their best sample higher than their poorest sample. The dull group, however, had a significantly higher percentage of "poorest" better than "best" than did the bright and average groups.

III. The relationship between perception and legibility in three groups differing in intelligence.

A. Performance on legibility was differentiated by factors of sex (girls consistently better writers), by intelligence groups (bright and average groups better writers than the dull); and by the conditions of the writing task (best samples were better than the fast).

B. The bright and average groups were consistently more accurate than the dull group in ordering their own handwriting samples according to legibility. The average group was more accurate than the bright, and girls were more accurate than the boys with the exception of the dull group where there were no differences. (The amount of confidence one can have in the sex difference is still questionable.)

C. In the Milwaukee study there was no clear-cut difference in the stability of the legibility of N^1 and N^2 samples (usual) of the bright, average, and dull groups. There was some suggestion that the average and bright groups and boys were more consistent in their normative perceptions than were the dull group and girls.

D. The bright and average groups and girls differed from the dull group and boys in the range of legibility produced in their

writing samples. Girls wrote over a greater range of quality than did boys.

E. There was a high correspondence between the high and average intelligence groups both on legibility and perception scores and evidenced a pattern which was different from the dull group.

F. There was no clear-cut difference between the three groups on perception scores but there is a significant one on legibility.

G. High perception scores do not predict either intelligence group or legibility of writing.

Discussion

The results of this study raise grave questions about the success of handwriting programs in providing the child with realistic models of handwriting performance which he can use to judge--and hence, improve--his handwriting performance. It might be expected by the end of the intermediate period that these children could effectively order handwriting specimens, of a size and slant to their own, in order of ease of reading. Yet few children in the study could do so with complete success. In addition, the children in the Madison study did not gain in this skill over the two years they were followed (1957-1959). There is every reason to assume that if this judgment were to improve, it would do so during the intermediate grades (grades 4-6). This perceptual development did not take place at a time when there would be every reason to assume that a child should be free from the initial task of learning the form of letters and words; would have moved in skill from drawing to efficient writing; and would now be in a period of development of normative and aspirational models which would form the basis for efficient handwriting skill.

It is very likely that in our traditional program of teaching writing, one kind of skill of forming and producing of letters and words has reached a plateau. To change this condition, skill training should not concentrate on letter

formation drill and writing routines but on the development and use of aspirational and normative models which would provide the references and dynamics for continued skill development.

The next research step to provide a sound basis for this new direction is obviously to try to establish and vary these normative and aspiration perceptions under different educational conditions. If this verifying research supports what is emerging here, the educational programs for much of the skill learning in the elementary school will need to be rewritten.

It is clear that, in general, bright and average children had perceptual patterns as defined by our study significantly different in all respects from those of the dull group. The only exception to this general statement is when a general perceptual score is contrasted with intelligence and legibility. Here it was shown that while intelligence and legibility were different for all groups, perception was not.

All groups experienced success in setting an aspirational level higher than their own level of writing and this aspirational model was more stable for all groups than was their normative perception of their own writing. There is some evidence to show that boys had a more consistent concept of their own writing than did girls, even though girls are better writers than boys.

This greater stability of the aspirational model was somewhat of a surprise although the ceiling effect of the scale may have artificially produced this effect for the bright students. Yet it is possible to see why aspirational models are and likely should be more fixed in a child's perceptual scheme than is his normative perception of his own and hopefully changing and improving handwriting. Since his own handwriting in this training period should be undergoing considerable modification, one should expect to get less consistency in the normative model and greater consistency in the aspirational model. What is disturbing is the fact that if the aspired to model is going to have any effect on the present

writing, the child will have to relate in some specific detail what the aspired to model means for modification of a perceived picture of the child's own handwriting. Unless this relationship can be implemented in consciously directed modification, these perceptual models can have no reality in fact.

It seems then, that the normative model for all groups is defective in several respects. Evidence for this is not only revealed in the research findings, but also in the observation by examiners of the great difficulty many children had in making the required judgments. We still do not know in detail why children perceive these handwriting tasks as they do, nor to what specific aspects of the task to which they are at times attending. The suggested next step in research points to a way for verifying the above emerging thesis regarding the directions in which to move in skill training.

The significant differences between bright and average groups and the dull do not suggest different programs of instruction in nature and conception but different programs in degree with suitable adaptations in materials and pacing of instruction. The results seem to suggest that if improved skill in handwriting is to be gained, similar improvements in the functioning of the aspiration and normative perceptual models would have to take place with dull children as well as with average and bright. Surely this is a critical problem in handwriting instruction; the present findings suggest directions to move in order to deal with it more adequately.

CHAPTER IV

PHYSIOLOGICAL AND MOTOR CORRELATES OF HANDWRITING LEGIBILITY

The Problem

The basic purpose of this study is to examine certain physiological and motor aspects of behavior in relation to handwriting legibility in bright, average, and slow learning children. The selection of legibility as a criterion for the examination of other correlative aspects of behavior provides a significant focus for examining the question of how good or poor handwriting is produced by these children. It may also provide a useful basis for certain recommendations for the improvement of handwriting instruction.

Two specific purposes have governed this investigation: (1) to observe changes in the affective state as measured by the galvanic skin response (GSR), muscular tension as measured by muscle action-potentials, and fine motor control as measured by point pressure variability in bright, dull, and normal children as the rate of writing conditions of handwriting are systematically varied; and (2) to consider the relationship between the legibility of handwriting produced under such conditions and the concomitant motor and physiological changes measured. In the succeeding chapter, the relationships of independent measures of motor control to handwriting performance are considered separately.

Since the present research is a continuation and extension of a series of studies, relevant aspects of the previous research conducted by the investigators will be summarized in the next section.

Previous Research by the Investigators

Prior to the initiation of the present study, two studies concerning the relationship between pressure in handwriting and legibility were completed, reported at meetings of the American Educational Research Association, and later published. The first of these was a comparative study of good and poor writers at the college level, the second an examination of the usual handwriting of randomly selected pupils in grades 4, 6, and 10. A third study reported upon very briefly at the February, 1959, meeting of A.E.R.A., was an exploratory attempt to record and study other psychophysiological variables than fine motor control in the writing act.

Pressure Patterns in Handwriting[1]

The first of these published studies, in 1955, (1) reviewed the paramount concern of most investigators with the relation of pressure phenomena in handwriting to personality characteristics; (2) developed a new method of interpretation of handwriting pressure patterns, the force variation ratio (FVR), and identified this as a measure of fine motor control related to handwriting legibility; (3) examined the effects of different rates of writing upon handwriting pressure patterns and legibility; and (4) investigated the inter-relationships among point pressure patterns, legibility and rate of writing. A standard sentence, "The quick brown fox jumps over the lazy dog," was used as the common writing task. A forced-choice technique for ranking all samples on a single continuum was employed as the legibility rating technique. Special instrumentation was employed to record point pressure in handwriting and to measure motor control.

---

[1]Theodore L. Harris and G. Lawrence Rarick, Pressure Patterns in Handwriting (Madison, Wisconsin: Committee for Research in Handwriting, Department of Education, University of Wisconsin, 1955). Later published as "Problem of Pressure in Handwriting," Journal of Experimental Education, XXVI (December, 1957), 151-78.

The 1955 study indicated no significant relationship between average point pressure and the variables of legibility and rate of writing. It was noted, however, that as an individual moved away from his normal rate of writing, his average word and sentence force increased. Moreover, a measure of variability in the application of force, the force variation ratio, was shown by rank-order correlation to be substantially related to average sentence pressure ($r = .69$) and to rate of writing ($r = .67$); and fairly highly related to legibility ratings ($r = .59$). A correlation of .54 was found between legibility and rate of writing. The relationships noted between the force variation data indicative of fine motor control and the variables of legibility and rate of writing suggested that further study of certain of these inter-relationships among children of elementary and high school age would be profitable. Accordingly such an investigation was next made.

Handwriting Pressure and Legibility in Children and Adolescents[2]

Data for the second study were gathered in 1955-56 in the field with the cooperation of two Wisconsin communities. This study differed from the earlier one in several important respects: (1) 144 subjects were randomly selected to include 12 boys and 12 girls at each of three grade levels--4, 6, and 10--in two communities; (2) only one sample of normal handwriting per subject was secured; and (3) legibility ratings of handwriting samples were made by means of a newly devised composite scale developed from samples drawn from grades 4, 6, and 10. In analyzing the force variation data it was shown that the force variation ratio for the two words, "fox jumps" was highly representative of that for the entire sentence. This finding reduced materially the time and effort involved in analysis of these and subsequent records.

---

[2]Theodore L. Harris and G. Lawrence Rarick, "Relationship Between Handwriting Pressure and Legibility of Handwriting in Children and Adolescents," Journal of Experimental Education, XXVIII (September, 1959), 65-84.

The legibility data were subjected to analysis of variance, the results showing significant sex and grade difference ($P < .01$), but none between communities. The sex difference favored girls whose improvement in legibility at the sixth and tenth grades offered a sharp contrast to the very slight progress shown by boys.

When the force variation data were treated by analysis of variance, significant differences were found for grade ($P = .01$), sex and community ($P = .05$). Force variation increased with grade level and was greater for girls than for boys. Thus an increase in force variation ratio accompanied an increase in legibility in the boys and girls of this study, a finding somewhat at variance with that of the earlier college study. Pearsonian correlations between force variation and legibility were uniformly low. Consistent grade differences but no sex differences were found for speed, which increased, and average pressure, which decreased.

Further analyses of the latter data indicated, however, that when the legibility and force variation ratio scores were examined as a percentage of the scores obtained at normal writing rates, high legibility was obtained on the average at slow writing rates and low legibility occurred when the writing rates were rapid. Likewise, when the writing rate was slow, variability in force was low and legibility was high; when the speed of writing was fast, variability in force was high and legibility was low.

An Exploratory Study of Affective State (GSR)

In the spring of 1957 data were gathered on a group of thirty third and fourth grade children representing five boys and five girls of bright, average, and slow learning ability at each grade level. One measure obtained on these children was the GSR as the children wrote a brief sentence "the fox jumps" successively in their usual, fast, and best manner. The GSR records were examined to see if any consistent pattern indicative of affective state could be ascertained.

A sensitive measuring apparatus, a Gilson Medical Electronics Polygraph, was used to record the psychophysiological activity. Since it was necessary to record continuous activity at relatively high speed to secure appropriate point pressure records, the records of GSR activity were not typical of those frequently secured in response to intense, intermittent stimuli. Accordingly, the procedure used in analysis was to measure off three segments of the GSR record: one equal to and parallel to the writing pressure recorded, a second segment of equal length preparatory to the actual writing, and a third segment of three second's duration following the actual writing. The GSR activity in each of these segments was examined with respect to the number, magnitude, and duration of "peak" activity from a baseline established during a rest period prior to writing. A sample record is shown in Appendix G.

Gross comparisons of GSR activity for the bright and average groups failed to reveal significant differences between them. A total pattern of GSR activity that is of considerable interest did, however, emerge. The records of subjects uniformly tended to show for normal, fast, and slow writing these characteristics: (1) the greatest number of peaks and the greatest average height of peaks in the preparatory phase--before the subject started writing; (2) a peak, usually of considerable duration, occurring within two seconds of the onset of writing and followed by a decline in the GSR; and (3) a sharp increase in the GSR within three seconds of the cessation of writing. Inspection of the fragmentary and incomplete writing records of the slow learning children suggested a similar pattern of affective activity.

Insofar as the GSR is an index of affect under autonomic control and hence involuntary, the uniform increase in GSR preparatory to writing suggests a mobilization of energy in anticipation of the task. The second rise in GSR shortly after the onset of writing may be related to effort expended in accomodation to the task, an adjustment which, however, appears to be readily made as the subject

engages in a familiar task and which in turn leads to a general decline in the affective state. The third and final GSR surge after the completion of the task appears to represent almost an expression of relief--sudden and sharp in incidence, and relatively quickly resolved. This pattern appeared to be characteristic not only of the performance of the bright and average group, but also by inspection of those records of the slow learning children whose writing records were on the whole fragmentary and incomplete.

Nature of the Variables Investigated

Although these earlier studies provided valuable although somewhat conflicting information on the relation of a measure of fine pressure variability and motor control (the force variation ratio--FVR) to the handwriting of children and adults and some preliminary data on related affective state, it became apparent that further understanding of handwriting behavior would come about only by probing deeper into its psychophysiological correlates. The production of letters and words is a mild psychomotor activity, but legible handwriting is exacting and requires precise movements of the arm, wrist, hand, and fingers. Similarly, it is an activity which may bring on general body tension and under certain circumstances may be mildly disturbing. While the earlier studies identified minute pressure changes as the letters were formed, no direct clues were available concerning the activity of the muscles controlling these movements. The preliminary attempt to measure affective overtones as conditions of writing were altered, was suggestive. It therefore appeared important to make a more comprehensive study of muscular response in the writing arm, generalized body tension, and autonomic responses indicative of the affective reactions of the individual, and to examine the relationships among these factors and that of legibility.

For these reasons, the scope of the research was extended to include data on the galvanic skin response (GSR) and on muscle action-potentials in selected motor units. The nature and significance of these variables is next described.

Affective State (GSR)

The GSR[3] (more properly called the galvanic skin reflex) is the most frequently used index of autonomic activity associated with affective states, particularly apprehension, arousal, and alertness. While the exact regulatory controls of the GSR are not clearly understood, it is well established that the nerve supply to the effector organ, the sweat glands, is exclusively sympathetic.[4,5,6] It is generally held that the electrical changes in the skin recorded as the GSR are due to alterations in membrane permeability associated with the secretory activity of these structures. Since the palms of the hands and the soles of the feet are liberally supplied with sweat glands, the GSR is most frequently picked up from one of these sources.

While there is not universal agreement concerning the meaning of the recorded GSR, there is considerable evidence that the response in animals and in man is an adaptive mechanism peculiarly sensitive to conditions of stress and anxiety.[7] It is on this premise that the GSR has enjoyed widespread use. However, differences of opinion exist in regard to the technique of recording the

---

[3]Wang draws a clear-cut distinction between the galvanic skin response and the galvanic skin reflex. The former, according to Wang, results from stimulation of the lumbar sympathetic chain and thus does not involve a sense organ or a sensory nerve. The galvanic skin reflex is a true reflex involving an affector organ and a sensory nerve. In this paper the term galvanic skin response is used synonymously with galvanic skin reflex.

[4]Donald B. Lindsley, "Emotion," Handbook of Experimental Psychology, ed. S. S. Stevens (New York: John Wiley and Sons, 1951), p. 475.

[5]Ging Hsi Wang, "The Galvanic Skin Reflex. A Review of Old and Recent Works from a Physiological Point of View," American Journal of Physical Medicine, 36 (October, 1957), Part I, 295-320.

[6]Ging Hsi Wang, "The Galvanic Skin Reflex. A Review of Old and Recent Works from a Physiological Point of View," American Journal of Physical Medicine, 37 (February, 1958), Part II, 35-57.

[7]C. W. Darrow, "Sensory, Secretory and Electrical Changes in the Skin Following Bodily Excitation," Journal of Experimental Psychology, 10 (1927), 197-226.

GSR and in the interpretations which are given to the records. In spite of this, there is general agreement on the types of stimuli which elicit the GSR. Most reports agree that GSR activity occurs in emotional states and that the strength of the response varies with the degree of anxiety.[8] Startle and apprehensive expectancy[9,10] and embarrassment and annoyance[11] bring on heightened GSR activity. Emotionally colored words also affect GSR.[12] There is some support for the belief that GSR reflects a preparatory mobilization of energy in relation to a stimulus situation.[13,14,15,16] Since a primary purpose of the current study was to systematically introduce changes in the writing task which conceivably might be disturbing to the subject, the GSR appeared to be an acceptable means for examining one aspect of the learner's anticipatory state and the effect of a mild motor activity upon subsequent affective responses.

---

[8]L. Welch, and J. F. Kubis, "Conditional PGR in States of Pathological Anxiety," Journal of Nucro-Mental Disorder, 105 (1947), 372-81.

[9]E. Patterson, "A Qualitative and Quantitative Study of the Emotion of Surprise," Psychology Monograph (1930), No. 181.

[10]E. Schiff, C. Dougan and L. Welch, "The Conditioned PGR and EEG as Indicators of Anxiety," Journal of Abnormal Social Psychology, 44, 549-552.

[11]R. S. Woodworth and H. Schlosberg, Experimental Psychology (New York: Henry Holt and Company, 1954), 137-59.

[12]Wang, "The Galvanic Skin Reflex...," Part I.

[13]C. W. Darrow, "The Galvanic Skin Reflex (Sweating) and Blood Pressure as Preparatory and Facilitative Functions," Psychological Bulletin, 33 (1936), 73-94.

[14]H. E. Jones, "The Galvanic Skin Reflex as Related to Overt Emotional Expression," American Journal of Psychology, 47 (1935), 241-51.

[15]H. E. Jones, "The Study of Patterns of Emotional Expression," Feelings and Emotion, The Mooseheart Symposium (M. L. E. Reymert),(New York: McGraw Hill, 1950), 161-68.

[16]Woodworth and Schlosberg, Experimental Psychology...

Muscular Tension (EMG)

It is well known that relaxed muscles are electrically silent, whereas muscles in a state of contraction produce minute electrical potentials which can be picked up by suitably placed electrodes and the signal amplified and recorded electronically. Electromyography offers a convenient means for studying the activity of muscles directly involved in coordinated movements as well as for recording muscular tension of a more general nature. The initiation and duration of muscular response can be accurately recorded by the EMG, but the meaning which can be attached to quantified action potential records is not always clear. While some hold that the EMG provides a faithful indication of developed muscular tension, others believe the recorded action potentials are a reflection of the conscious effort expended rather than the actual tension developed. For example, Gregg and Jarrard,[17] Inman,[18] Knowlton,[19] Lippold,[20] and Small[21] report that the relative magnitude of the surface electromyogram and the force of the muscular contraction are approximately linear. On the other hand, Eason[22] found that the amplitude of action potentials increased during fatigue. By and large the preponderance of evidence indicates that within reasonable limits the EMG is a function of the tension developed until the muscles show marked fatigue.

---

[17]Lee W. Gregg and Leonard E. Jarrard, "Changes in Muscle Action Potentials During Prolonged Work," Journal of Comparative and Physiological Psychology, 51 (October, 1958), 532-35.

[18]V. T. Inman, et al, "Relation of Human Electromyogram to Muscular Tension," Electromyography and Clinical Neurophysiology, 4 (1952), 187-94.

[19]G. C. Knowlton, et al, "Electromyography of Fatigue," Archives of Physical Medicine, 32 (1951), 648-52.

[20]O. C. J. Lippold, "The Relation Between Integrated Muscle Action Potentials in Human Muscle and Its Isometric Tension," Journal of Psychology, 117 (1952), 492-99.

[21]A. M. Small, Jr. and N. B. Gross, "Integrated Muscle Action Potentials in a Weight-Lifting Task as a Function of Weight and Rate of Lifting," Journal of Comparative and Physiological Psychology, 51 (1958), 227-29.

[22] R. G. Eason, An Electromyographic Study of Impairment and Estimates of Subjective Effort Associated with Voluntary Muscular Contraction (San Diego, California: U. S. Navy Electronics Laboratory, May, 1959), Research and Development Report Number 898, 27 p.

Studies pertaining to psychomotor performance of a milder nature suggest that specific patterns of EMG activity become characteristic of the task.[23,24,25,26] Furthermore, there is some evidence to indicate that progressive isolation of the segment directly involved in the activity occurs as proficiency in the task developed.[27] However, where learning of psychomotor skills is occurring under conditions of frustration or high motivation, widespread generalized muscular tension is frequently observed.[28,29,30,31]

It seems particularly important that an investigation of the handwriting of children calling for a systematic variation in writing tasks should provide information on the electromyographical changes in the muscles directly associated with writing and should also give data on the general body tension as indicated by electrical changes in more remote regions of the body.

---

[23] R. S. Daniel, "The Distribution of Muscular Action Potentials During Maze Learning," Journal of Experimental Psychology, 24 (1939), 621-29.

[24] G. L. Freeman, "Spread of Neuro-Muscular Activity During Mental Work," Journal of Genetic Psychology, 5 (1931), 479-94.

[25] R. L. Henderson, "Remote Action Potentials at the Moment of Response in a Simple Reaction Time Situation," Journal of Experimental Psychology, 44 (1952), 238-41.

[26] C. R. Oldreyd and B. Moskowitz, "Optimal Tensions for Various Types of Tasks as a Function of Proximity to Focal Musculature," Proc. Oklahoma Acadamy of Science, 33 (1952), 265 p.

[27] Freeman, "Spread of Neuro-Muscular..."

[28] R. C. Davis, "The Relation of Muscle Action Potentials to Difficulty and Frustration," Journal of Experimental Psychology, 23 (1938), 141-58.

[29] H. D. Meyer, "Reaction Time as Related to Tensions in Muscles Not Essential in the Reaction," Journal of Experimental Psychology, 39 (1949), 96-113.

[30] A. A. Smith, "An Electromyographic Study of Tension in Interrupted and Completed Tasks," Journal of Experimental Psychology, 46 (1953), 32-36.

[31] J. B. Stroud, "The Role of Muscular Tension in Stylus Maze Learning," Journal of Experimental Psychology, 14 (1931), 606-31.

Fine Motor Control (Force Variation)

Since handwriting is a perceptual-motor activity which presumably requires a high degree of fine motor control in the application of force during the handwriting act, a measure of the latter might usefully serve as an index of the degree of such motor control. Such a conception of motor control in handwriting refers to a vertical dimension of control exerted in the course of writing, rather than to a horizontal dimension of motor control such as might be inferred from handwriting legibility or from specially constructed tasks designed to measure speed and accuracy of motor control independent of the handwriting task. Findings in the latter dimension are discussed in Chapter V.

The force variation ratio (FVR), a ratio in which the cumulative amount of change in pressure was divided by time (in seconds), was used to measure fine motor control. The use of such a ratio made possible comparisons between samples differing in speed and average point pressure.

The Population and Methodology of the Study

Population

In the spring of 1959, 30 children of approximately the chronological age of 135 months and drawn from the fifth and sixth grade levels of the Madison Public Schools were brought to the Basic Skills Laboratory of the Department of Education for testing. These children comprised three groups of ten children, five girls and five boys each, randomly drawn within sharply differentiated levels of intellectual ability from the school population. Average I.Q.'s on the Wechsler Intelligence Scale for Children for the three groups were: Bright - 120; Average - 100; and Dull - 71. With five exceptions in the slow learning group, these were the same children seen in the laboratory two years previously at which time exploratory data of a more limited nature were gathered. A complete description of this sample is found in Chapter III.

Instrumentation and Testing Conditions

All tests were conducted in a small electrically shielded booth 6' x 6' x 6'. The booth was large enough to provide space for the adjustable handwriting table, the subject, and one investigator who supervised the experimental tasks. A Gilson Eight Channel Polygraph placed just outside the testing booth was used to provide a continuous record of the psychomotor and physiological changes. EMG and GSR leads from the subject and leads from the strain gauge pressure recording system ran directly to the electronic recorder. The placement of the instrumentation in relation to the booth is shown on the accompanying diagram.

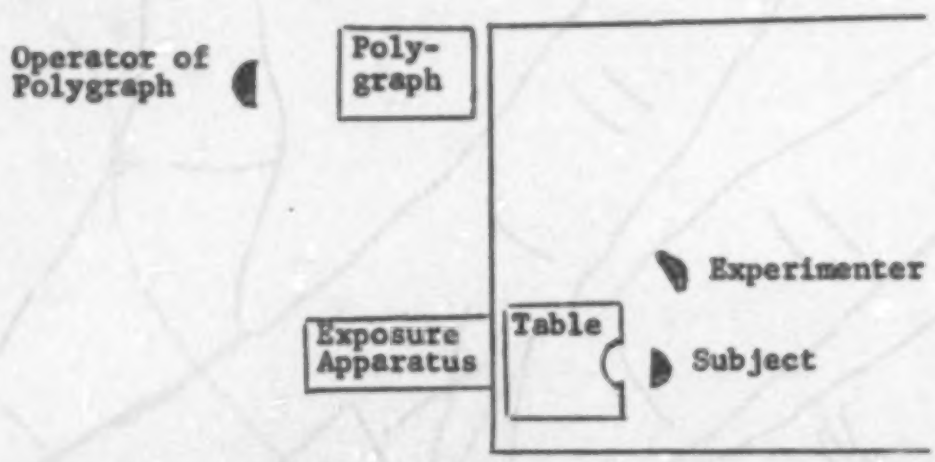


Fig. 3. Diagram of Testing Booth and Placement of Apparatus and Personnel

Silver disk surface electrodes, eight mm. in diameter, containing electrode paste and securely anchored with adhesive tape, provided the contacts for the electromyography. The electrodes for securing action potentials immediately associated with writing were mounted over the extensor communis digitorum on the dorsal surface of the forearm approximately two inches below the elbow. Surface electrodes mounted over the upper trapezius were used to secure action potentials of muscular tension in the shoulder-neck region. Specially designed dual metallic electrodes were anchored to the palm of the left hand giving contact for the GSR. (See Fig. 4.) A metal platen mounted flush with the top of the writing table and resting on strain gauges (See Fig. 5) was used to detect

Fig. 4. Placement of Dual Metallic Electrodes for Contact for the GSR.

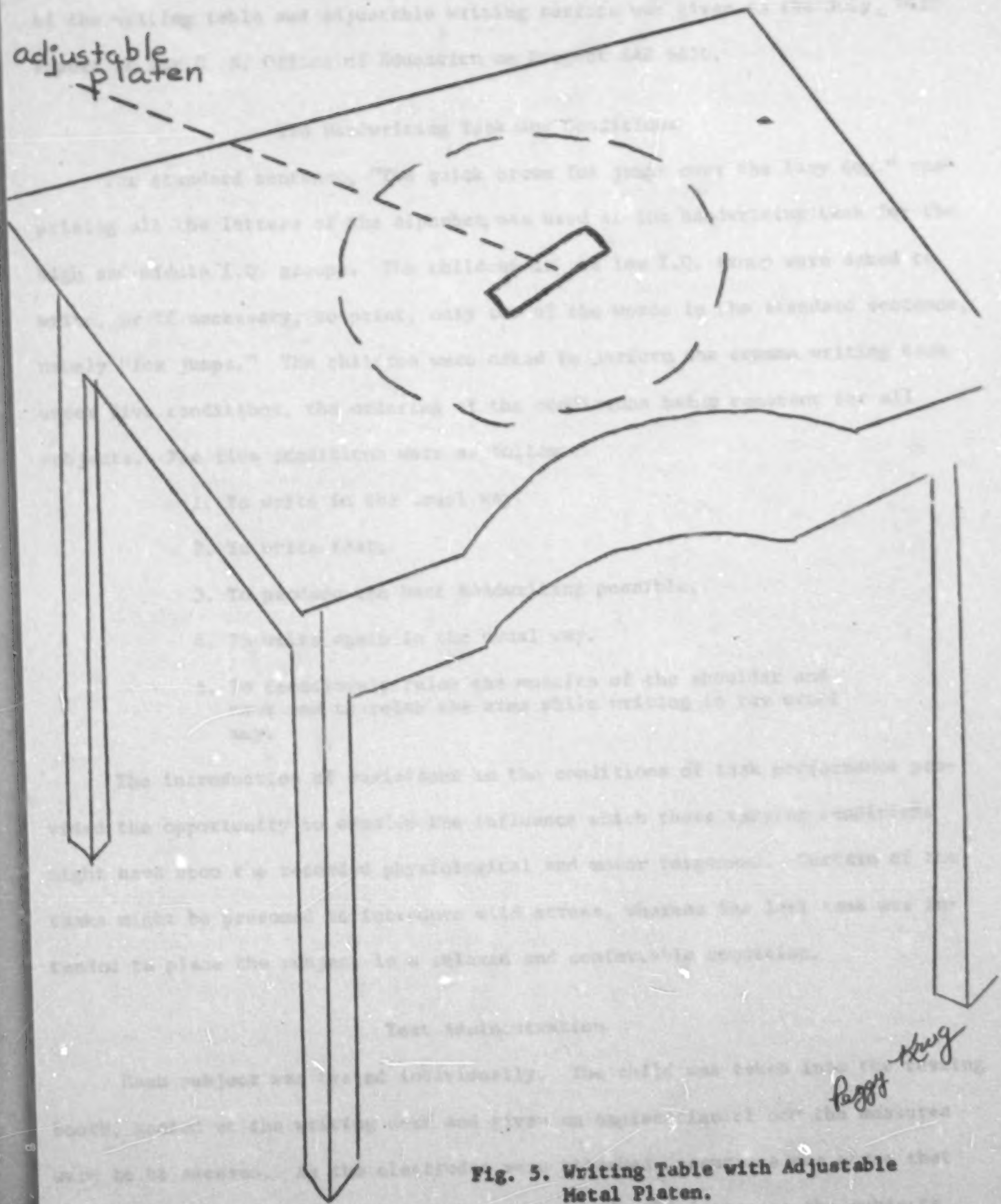


Fig. 5. Writing Table with Adjustable Metal Platen.

fluctuations in point pressure during the writing act. A complete description of the writing table and adjustable writing surface was given in the July, 1957 report to the U. S. Office of Education on Project SAE 6436.

The Handwriting Task and Conditions

The standard sentence, "The quick brown fox jumps over the lazy dog." comprising all the letters of the alphabet, was used as the handwriting task for the high and middle I.Q. groups. The children in the low I.Q. group were asked to write, or if necessary, to print, only two of the words in the standard sentence, namely "fox jumps." The children were asked to perform the common writing task under five conditions, the ordering of the conditions being constant for all subjects. The five conditions were as follows:

1. To write in the usual way.
2. To write fast.
3. To produce the best handwriting possible.
4. To write again in the usual way.
5. To consciously relax the muscles of the shoulder and neck and to relax the arms while writing in the usual way.

The introduction of variations in the conditions of task performance provided the opportunity to examine the influence which these varying conditions might have upon the recorded physiological and motor responses. Certain of the tasks might be presumed to introduce mild stress, whereas the last task was intended to place the subject in a relaxed and comfortable condition.

Test Administration

Each subject was tested individually. The child was taken into the testing booth, seated at the writing desk and given an explanation of how the measures were to be secured. As the electrodes were attached, assurance was given that no discomfort would result from the recording. In all instances the subjects

appeared to be perfectly at ease. Directions were given concerning the writing tasks and the cues which one investigator would give during the experiment. The operation of the recorder was handled by the second investigator stationed outside the booth so that distractions were minimal. The following sequence was followed:

1. During the preparatory period the subject was asked to relax until the GSR became stable. The recorder was started early in this period.
2. The subject was asked to tense his right fist. This provided the EMG record of high tension of the muscles of the arm and also gave evidence of spread of tension which was picked up from the shoulder-neck electrodes.
3. The subject was asked to rest.
4. The subject was then directed to firmly tense the muscles of the shoulder-neck region. (All subjects were able to consciously tense the muscles of this region, usually with tension spreading to the writing arm.)
5. The subject was given a short rest.
6. Next, the subject was asked to write the standard sentence in his usual way.
7. A rest period followed.
8. The handwriting task was repeated with the subject directed to write fast.
9. A rest period followed.
10. The subject was next asked to repeat the task but to write his very best.
11. A rest period was given.
12. The task was repeated with the subject writing in his usual manner.
13. The subject was given a rest.
14. The subject was asked to tense his shoulder and neck firmly so that he could sense the feeling of tension and thus be better able to consciously relax his muscles.
15. A rest period was given.
16. The subject was then directed to write in his usual manner, but was asked to keep his shoulder and neck as loose and relaxed as possible.

17. A final rest period was given and the testing ended when the GSR returned to normal.

Method of Processing the Data

The simultaneous recording of the GSR, EMG's and point pressure gave a permanent and synchronized record of these responses as shown in Appendix G. Thus, fluctuations in one measure over a period of time could be compared with fluctuations in other recorded responses. Likewise, changes in a given response could be observed as the conditions of the task were changed.

The variations in point pressure (FVR) were measured by the use of a map reader. A ratio was computed by dividing the length of the force curve (from the initial peak to the final peak prior to return to the base line) by the length of the base line subtended by the initial and final force peaks. By extending perpendicular lines from the point of origin and from the point of termination of the pressure record through the GSR and EMG tracings, the writing segment of the record was clearly defined for all measures. Since the paper speed was constant, writing time could be accurately determined.

It should be noted that the characteristic rise in GSR activity in the pre-writing, writing, and post-writing periods which was noted in the earlier exploratory study was likewise found in most records in this study. However, in the following analysis, measurements were confined solely to phenomena occurring <u>within</u> each of the five writing tasks as defined by the duration of the handwriting pressure record.

The GSR curves varied from subject to subject and from task to task. In some instances the deflection continued throughout the writing task, whereas in other cases there occurred a series of deflections with the curve reaching the base line at several points. The score which seemed to be most appropriate for our purposes was determined by measuring the total area under the GSR curve for

a given writing task with a planimeter and by dividing this by the length of the base line, thus giving the average deflection of the GSR.

The quantification of the EMG's was done by the use of a specially designed transparent plastic scale which was superimposed over the EMG records. The scale, three centimeters wide, was graduated by parallel lines two millimeters apart with a zero line which was placed over the midline of the EMG tracing. In taking measurements, the scale was placed over a segment of the EMG tracing with the zero line of the scale at the mid-point of the tracing. In quantifying the EMG tracings, the first digit of the EMG score was taken as the scale value farthest from the zero line which encompassed a minimum of 10 deflections in the area bounded by the scale. The next digit was taken as the number of deflections in the next higher scale value and so on until all deflections were accounted for. Thus, a score of 761 would result when the seventh scale unit was the one farthest removed from the center line containing a minimum of 10 deflections. The last two digits, 6 and 1, indicated that 6 deflections appeared in scale value eight and one deflection appeared in scale value nine. For each writing task, three EMG segments were chosen for measurement, one near the beginning, one near the middle, and one near the end of the record. Thus segments corresponding to the words "quick," "jumps," and "lazy" were selected for measurement and the average score based upon these three segments constituted the EMG score for that task.

Results of the Study

Writing Speed of Average and Bright Children

While a major purpose of the investigation was to examine the recorded variations in the four psychophysiological responses as the conditions of the writing task were changed, it seemed important that consideration be given to individual and group writing rates by conditions as an aid in interpreting the

recorded responses. Since one of the specific task conditions called for fast writing, it was important to know if the request for the fast sample was actually the child's fastest writing. Likewise, in the task condition calling for the child's best writing, one might hypothesize that the child would tend to write slowly since it has been shown that speed of writing is an important factor in legibility. The investigators were also especially interested in observing writing rates of those who completed the standard sample; in this case, the bright and average boys and girls.

Means and standard deviations of writing times for bright and average boys and girls for the five task conditions are given in Table 43. It will be noted

TABLE 43

MEANS AND STANDARD DEVIATIONS OF WRITING TIMES*
FOR BRIGHT CHILDREN AND CHILDREN OF AVERAGE
INTELLIGENCE ACCORDING TO CONDITION

Writing Task	Bright I.Q. Children				Average I.Q. Children			
	Boys		Girls		Boys		Girls	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Normal[1]	49.48	16.44	33.82	4.68	35.26	3.29	33.03	7.97
Fast	31.72	10.22	23.16	2.41	25.68	3.19	23.92	5.06
Best	51.26	16.65	33.78	4.96	38.78	7.01	36.10	7.07
Normal[2]	39.22	11.30	29.36	2.13	29.10	3.62	30.86	8.24
Relaxed	39.70	9.53	30.17	3.85	32.32	3.80	32.28	8.18

*In seconds.

that for all groups the mean writing time was lowest for the conditions requesting fast writing. Inspection of individual writing times shows that the requested fast sample was the fastest of the five samples in every case. It should also be noted that for three of the four groups, the subject's intended best sample was produced on the average at the slowest rate. The writing times by conditions of the bright girls and of the boys and girls of average intelligence were similar, but the bright boys wrote considerably slower and showed less homogeneity in their writing times. Examination of individual scores showed that twelve of

the 20 subjects wrote at their slowest rate when attempting to produce their best handwriting. The child's concern about legibility would therefore appear to be an important factor in keeping speed of writing down.

The speed data for four of the five writing conditions (omitting the second normal sample) were subjected to an analysis of variance with the results shown in Table 44. Significant differences in speed (P <.01) were found only among the

TABLE 44

SUMMARY OF ANALYSIS OF VARIANCE FOR WRITING RATES IN BRIGHT AND AVERAGE I.Q. CHILDREN

	df	Sum Sq.	Mean Sq.	F
I.Q.	1	406.81	406.81	1.400
Sex	1	1035.37	1035.37	3.564
I.Q. x Sex	1	618.26	618.26	2.128
Sub x Cell	16	4648.15	290.51	16.648**
Conditions	3	2243.99	748.00	42.865**
Con/I.Q.	3	81.89	27.97	1.602
Con/Sex	3	107.80	35.93	2.059
Con/Sex/I.Q.	3	52.18	17.39	--
Con/Sub/Cells	48	837.73	17.45	--

**Significant at the .01 level.

writing conditions, not for sex or intelligence. While the F ratios for sex and for conditions by sex were not statistically significant (P <.05), a review of the mean writing times (see Table 43) shows that the bright boys wrote at a substantially slower rate than the other three groups under each condition due primarily to two very slow writers. Inspection of this table shows little difference in mean writing times for the other three groups under any of the writing conditions.

In order to determine the differential effects of the task conditions upon individual writing rates, intercorrelations for writing times by conditions were computed for boys and girls separately and are shown in Table 45.

These intercorrelations show that there is high individual consistency in writing speed for the five conditions of writing. In other words, the fast

TABLE 45

INTERCORRELATIONS AMONG WRITING RATES BY TASK CONDITIONS FOR BRIGHT AND AVERAGE BOYS AND GIRLS

Conditions	Boys					Girls				
	N^1	F	B	N^2	R	N^1	F	B	N^2	R
N^1	--	.787	.936	.937	.941	--	.923	.795	.863	.936
F		---	.741	.929	.906		---	.719	.876	.870
B			---	.858	.927			---	.774	.879
N^2				---	.965				---	.927
R					---					---

writers tend to maintain relatively fast writing rates regardless of condition of writing while the reverse is true for slow writers. It would appear that the individual, while producing handwriting at different rates of writing, nevertheless develops a characteristic writing tempo which governs his relative rates of writing. This may suggest that a set for speed is operating as a significant factor in the handwriting act.

Fine Motor Control (FVR) of Average and Bright Children

The earlier studies had suggested that the FVR is a reflection of one aspect of a motor set in handwriting which may be reasonably stable within the individual, but which may be readily disturbed as the individual moves away from his habitual manner of writing. Changes to a faster than usual writing rate, for example, may then disturb the motor set and thus tend to reduce legibility and increase the FVR. It should be kept in mind that concepts as to what constitutes fast writing differ widely among children, so there is no way of being certain how much the request to write fast really pushes the child. Furthermore, some children undoubtedly have the potential for increasing their writing speed considerably without strain, whereas for others a modest increase is accomplished with considerable disturbance of the motor set. The production of two usual samples with an intervening fast sample tends to be accompanied by an increase in the writing rate in the second usual sample, as reference to the mean

writing times in Table 43 shows. This occurs even though the slowest writing time for their best sample also intervened between the two usual samples.

In the present study it was hypothesized that if the various conditions under which the child wrote the standard sentence were sufficiently different, such differences might be reflected in the FVR. Mean FVR's, obtainable only for children of bright and average intelligence, are shown in Table 46. While the

TABLE 46

MEANS AND STANDARD DEVIATIONS OF FORCE VARIATION RATIOS,* AN INDEX OF FINE MOTOR CONTROL, FOR BRIGHT AND AVERAGE I.Q. CHILDREN ACCORDING TO TASK CONDITION

Writing Task	Bright I.Q. Children				Average I.Q. Children			
	Boys		Girls		Boys		Girls	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Normal[1]	1.66	.42	1.83	.28	1.92	.20	1.68	.60
Fast	1.88	.62	2.51	.45	2.32	.72	1.90	.40
Best	1.85	.50	2.15	.61	1.60	.20	1.43	.44
Normal[2]	1.76	.49	2.34	.79	1.82	.58	1.47	.37
Relaxed	1.42	.28	1.66	.42	1.62	.45	1.22	.36

*The higher the FVR, the less the degree of fine motor control, and vice versa.

differences in the means within groups for the various writing conditions are not great, the FVR for fast writing was the highest for each group. There also was a tendency for the mean FVR's to be low in relaxed writing. In 12 of the 20 cases the highest FVR occurred as fast samples were produced and in 12 cases the lowest FVR occurred with relaxed writing. While these data are in no sense conclusive, they do lend support to the hypothesis that fast writing adversely affects motor control as measured by the FVR and that relaxed writing appears to be an asset in steady application of force. It will also be noted that the highest average FVR in each of the four groups occurred under the fast writing condition.

An analysis of variance for the fine motor control (FVR) data is shown in Table 47. It is evident that for these children the force variation ratio is

TABLE 47

SUMMARY OF ANALYSIS OF VARIANCE FOR FINE MOTOR CONTROL (FVR) IN BRIGHT AND AVERAGE CHILDREN

	df	Sum Sq.	Mean Sq.	F
I.Q.	1	.4899	.4899	--
Sex	1	.0031	.0031	--
I.Q. x Sex	1	2.0288	2.0288	2.877
Sub/Cell	16	11.2795	.7050	6.184**
Conditions	3	4.5300	1.5100	13.245**
Con/I.Q.	3	.7804	.2601	2.281
Con/Sex	3	.1054	.3513	3.081*
Con/Sex/I.Q.	3	.3136	.1045	--
Con/Sub/Cells	48	5.4699	.1140	

*Significant at the .05 level.
**Significant at the .01 level.

significantly affected by the writing condition (P = .01). While the analysis showed no overall sex difference in FVR, the significant (P < .05) interaction of conditions by sex indicated that condition had a differential effect on the FVR of the two sexes. It may be inferred from this evidence of interaction between sex and conditions of writing that the boys and girls of this sample differed in their force variation ratio as the condition of writing was varied. Reference to Table 46 indeed indicates that as boys and girls move from one condition of writing to another, the differences in the force variation ratio are not consistently in the same direction.

Intercorrelations among the force variation ratios for the five conditions of writing were obtained for each sex and are given in Table 48.

TABLE 48

INTERCORRELATIONS AMONG FORCE VARIATION RATIOS BY TASK CONDITIONS FOR BRIGHT AND AVERAGE BOYS AND GIRLS

Conditions	Boys					Girls				
	N^1	F	B	N^2	R	N^1	F	B	N^2	R
N^1	--	.328	.510	.334	.784	--	.428	.607	.454	.845
F		---	.882	.887	.494		---	.851	.887	.529
B			---	.948	.733			---	.941	.741
N^2				---	.613				---	.641
R					---					---

It will be noted that all correlations are positive, ranging from .328 to .948 for the boys, and from .428 to .941 for the girls. The pattern of the intercorrelations is strikingly similar for both sexes, the highest of these involving the intercorrelations among the fast, the best and the second normal writing conditions. These high correlations suggest that the three middle tasks had little differential effect upon the individual's fine motor control. On the other hand, differential effects of conditions are evidenced by the lower correlation involving the first and last task conditions in the series; namely, the first normal and the relaxed writing conditions. It is conceivable that the adaptation to the initial task and the orientation to the novel nature of the relaxed writing task made these circumstances less appropriate for revealing the individual's motor set than did the other conditions of writing.

Within Subject Agreement Between Conditions and Between Psychophysiological Factors

Within vs. Between-Subject Comparisons

The method of quantifying the FVR and EMG records was described earlier. It should be kept in mind that GSR recordings are highly unstable and that wide variations in the GSR response occur among individuals and for a given individual at different times. Therefore, a measured GSR response when used for inter-individual comparison is likely to be meaningless. While GSR responses for a given individual at specified points in time provide useful information on the direction of change for that individual, the measured increments of change for one individual may not be valid for making comparisons of GSR changes among individuals. In the case of EMG records, differences in skin conductivity pose a similar problem as do electrode placement and the adequacy of contact established. Therefore, it did not seem justifiable to use raw scores for either within or between individual comparisons. Hence, each individual's five raw scores obtained under each of the writing conditions for a specific psychophysiological

measure were placed in rank order, from greatest to least activity. Each subject thus had four sets of five ranked scores, one set for each of the four psychophysiological variables (GSR, FVR, EMG-arm, and EMG-shoulder) for each of the five writing tasks.

While the exact magnitude of the response was not used as such in this type of analysis, questions of "greater than" or "less than" could be explored. Since a primary concern of this study was change of response pattern and not magnitude of response this technique was thought to be acceptable to explore differences in these measures by task and subject.

Within Subject Agreement by Conditions

It might be hypothesized that the handwriting tasks which the children were asked to perform were so similar that the ordering of the psychophysiolgical responses by conditions might be a matter of chance. In other words, under chance conditions the frequency of cases in each rank for each of the five writing conditions should be approximately the same. The distribution of ranks according to the five handwriting conditions grouped by I.Q. level is given in Table 49. A chance distribution would result in an N of 8 in each cell since

TABLE 49

FREQUENCY TABULATION OF RANKS FOR EACH OF THE WRITING TASKS ACCORDING TO I.Q. GROUPS

Writing Task	High IQ Group (N=10) Rank					Ave. IQ Group (N=10) Rank					Low IQ Group (N=10) Rank				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Normal[1]	6	12	10	8	4	5	16	10	6	2	6	5	5	10	4
Fast	23	10	4	3	0	23	7	6	0	3	14	7	7	2	0
Best	5	5	10	14	4	5	8	7	15	5	5	8	10	3	3
Normal[2]	5	10	10	9	6	5	7	9	13	4	2	8	5	12	3
Relaxed	1	5	7	9	17	2	6	8	5	18	4	2	3	4	17

four records were obtained on each subject under each writing condition. By visual inspection it is clear that under fast writing conditions the number of

cases in all three groups with a rank of 1 is high and the number with a rank of 5 is low. Conversely, under conditions of relaxed writing, a disproportionate number of cases fall in rank 5 with only a few appearing in rank 1. Without attempting at this point to differentiate ranks according to psychophysiological measures, the data suggest that fast writing tends to maximize psychophysiological reactions whereas relaxed writing appears to have the opposite effect. Under the condition requesting the best writing, the highest frequency for middle and high I.Q. groups was rank 4, indicating that for a sizable number of subjects the incidence of measured reactions tended to be relatively low.

Within Subject Agreement Among Factors by Conditions

To examine the question of within subject agreement among the measures for the five writing conditions, the Kendall Coefficient of Concordance for ranked data was employed. A Kendall-W was computed for each of the 30 boys and girls representing all three levels of intelligence. The results are given in Table 50. It will be noted that all coefficients are positive and that 10 of the 20

TABLE 50

WITHIN SUBJECT AGREEMENT AMONG THE FOUR PSYCHOPHYSIOLOGICAL MEASURES FOR THE FIVE CONDITIONS OF WRITING AS DETERMINED BY THE KENDALL COEFFICIENT OF CONCORDANCE: W†

Bright Children (4 variables)		Children of Av. Intelligence (4 variables)		Slow Learning Children*** (3 variables)		
Boys	Girls	Boys	Girls	Boys	Girls	N=30
.671*	.570*	.134	.687**	.675	.238	6
.133	.822*	.225	.725**	.800*	.622	6
.223	.325	.550	.650*	.266	.111	6
.462	.559*	.553*	.275	.462	.422	6
.625*	.578*	.194	.384	.462	.377	6

*Significant at .05 level.
**Significant at .01 level.
***GSR, arm EMG, and shoulder EMG only usable measures.
† Each coefficient represents the correlation of these factors for a given subject.

in the middle and high I.Q. groups are significant at or beyond the .05 level.

Since it was not possible to obtain comparable FVR records for most of the slow

learning children, the coefficient of concordance for these children was computed on only three variables. It is apparent that even moderately high agreement among the ranks of these variables for the slow learning children was highly unusual. On the other hand, seven of the 10 girls of average and high intelligence showed coefficients of concordance significantly greater than chance, whereas this was true of only three boys. Apparently for girls, who it will be shown are better writers as a group than are the boys, these functions tend to fluctuate similarly according to the task condition. It might be hypothesized that these psychophysiological mechanisms tend to reinforce each other, certain writing conditions bringing about relatively high responses while others yield relatively low responses. This might suggest greater ability on the part of the girls to coordinate their response mechanisms including those which are only indirectly related to the handwriting act.

Analysis of the Legibility of the Samples

Thus far the analysis of data has been confined to the psychophysiological measures employed in the study. However, the development of improved scaling techniques for the determination of legibility scores enabled comparisons to be made to include this important factor.

The legibility scores of the bright and average children were subjected to analysis of variance with the results shown in Table 51. That legibility is significantly affected by the condition of writing appears clearly established ($P < .01$). Slightly less significant differences ($P < .05$) are obtained for sex. The direction of mean sex differences, as shown in Table 52, in every condition of writing favors the girls who systematically tended to produce the more legible handwriting. Although no significant differences in legibility were found for I.Q., it will be noted in this table that the means of the legibility scores grouped according to I.Q. and sex are higher for the high I.Q. children than for average I.Q. children. These differences are more marked for boys than for

TABLE 51

SUMMARY OF ANALYSIS OF VARIANCE FOR LEGIBILITY FOR BRIGHT AND AVERAGE CHILDREN

	df	Sum Sq.	Mean Sq.	F
I.Q.	1	.7354	.7354	1.181
Sex	1	4.6706	4.6706	7.498*
I.Q. x Sex	1	.2703	.2703	--
Sub/Cells	16	9.9671	.6229	17.797**
Conditions	3	.8949	.2983	8.523**
Con/I.Q.	3	.0202	.0067	
Con/Sex	3	.0536	.0179	
Con/Sex/I.Q.	3	.0053	.0018	
Con/Sex/Cells	48	1.6783	.0350	

**Significant at the .01 level.
*Significant at the .05 level.

TABLE 52

MEANS AND STANDARD DEVIATIONS OF LEGIBILITY SCORES OF BRIGHT CHILDREN AND CHILDREN OF AVERAGE INTELLIGENCE FOR FOUR CONDITIONS

	Bright I.Q. Children				Average I.Q. Children			
	Boys		Girls		Boys		Girls	
Writing Task	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Normal	2.55	.26	2.96	.33	2.25	.63	2.86	.35
Fast	2.32	.11	2.66	.30	2.06	.36	2.61	.28
Best	2.63	.33	2.92	.36	2.32	.46	2.88	.49
Relaxed	2.48	.24	2.90	.37	2.11	.56	2.80	.30

girls. Legibility likewise appears to vary inversely to a fairly substantial degree with speed of writing, as a rank order correlation of -.55 between writing time and legibility on normal samples indicates; i.e., high legibility tends to be associated with slow speed, and vice versa.

Intercorrelations among the legibility ratings for the five conditions of writing for each sex are given in Table 53. The correlations are all positive and relatively high, especially among the boys whose range of scores was considerably greater than that of the girls. These findings indicate marked individual stability in legibility regardless of the condition as was shown to be true for writing speed and to a lesser extent for fine motor control.

TABLE 53

INTERCORRELATIONS AMONG LEGIBILITY RATINGS BY TASK CONDITIONS FOR BRIGHT AND AVERAGE BOYS AND GIRLS

Conditions	Boys					Girls				
	N^1	F	B	N^2	R	N^1	F	B	N^2	R
N^1	--	.917	.901	.819	.856	--	.892	.951	.703	.774
F		---	.932	.833	.817	--	---	.861	.654	.840
B			---	.827	.771			---	.632	.722
N^2				---	.828				---	.682
R					---					---

In other words, such differences as may have been produced by task conditions affected changes among these factors in individuals in a similar direction and to a somewhat similar extent. As might be thus expected, legibility tended to correspond to only a very slight degree with speed and with the force variation ratio, regardless of writing condition.

Interestingly enough, analysis of the normal, fast, best, and relaxed samples of each of the subjects in the bright and average groups revealed that the most legible sample for the individual was not always produced under the condition of best writing. In point of fact, only seven of the 20 subjects produced their most legible sample under the best writing condition, while five more did so under the relaxed condition. Seven wrote their most legible sample under the first normal condition, and one, a very poor male writer, actually wrote his best sample under the fast condition. Thus while the group data show some consistency in legibility trends as a function of condition of writing, an individual's most legible handwriting sample is not necessarily dictated by the condition of his writing.

While the analyses thus far show a substantial tendency for legibility to decline when subjects are asked to write fast, a comparison of legibility and speed data for individual cases is enlightening. For example, a boy in the high group wrote at the slowest rate among the 20 subjects under all four writing conditions, yet his average legibility score for the four conditions gave him a rank

of seventeen in legibility. It is interesting to note that his writing time for the four tasks ranged from 52.1 seconds to 74.0 seconds while his legibility scores were closely bunched, ranging from 2.21 to 2.45. It is also worthy of note that the two poorest writers, both boys of average I.Q., produced samples at a slower rate under each condition than the mean writing time for their sex and I.Q. group. On the other hand, the girl in the bright group who produced the sample of highest legibility under all four writing conditions ranked 13th in speed, whereas the boy in the average group who consistently produced fast samples and had the lowest average writing time was ranked fourth in legibility. Thus, it is evident that absolute writing time was not necessarily the critical factor in the production of legible handwriting in this experiment, for good handwriting was produced at rapid rates and poor handwriting was produced at slow rates. The evidence suggests that writing as it affects legibility is a relative matter for the individual and that writing which is substantially faster than the usual rate is likely to result in a decline in legibility.

Legibility was also examined by conditions of writing in relation to fine motor control (FVR) to determine the number and per cent of cases in which the legibility and force variability was greater than that for the normal sample. The results are shown in Table 54. Only half of the 20 subjects produced a best

TABLE 54

NUMBER AND PER CENT OF BRIGHT AND AVERAGE SUBJECTS PRODUCING SAMPLES WITH LEGIBILITY AND FORCE VARIATION GREATER THAN THAT OF THE FIRST NORMAL SAMPLE

	N^1		F		B		N^2		R	
	N	%	N	%	N	%	N	%	N	%
Legibility	-----		1	5	10	50	7	35	8	40
FVR	-----		15	75	6	30	5	25	8	40

sample that was better than their first normal, while less than half did so under each of the last two conditions of writing. Force variability increased in

75 per cent of the subjects during fast writing but tended to decrease rather than increase for most subjects in the latter three conditions of writing.

If it is held that a decrease in force variability should tend to accompany an increase in legibility, the figures in Table 55 appear to lend some support

TABLE 55

NUMBER AND PER CENT OF BRIGHT AND AVERAGE SUBJECTS IN WHICH AN INCREASE IN LEGIBILITY IS ACCOMPANIED BY A DECREASE IN FORCE VARIATION

N^1		F		B		N^2		R	
N	%	N	%	N	%	N	%	N	%
----		16	80	10	50	12	60	11	55

to this assumption. In 50 per cent or better of the cases for all conditions subsequent to the first normal, an increase in legibility is accompanied by a decrease in the FVR, and vice versa. This is most noticeable for the conditions of fast writing in which 80 per cent of the subjects acted in the predicted direction. Thus, while the per cent of agreement is not great, the substance of the hypothesis is confirmed for more cases than not and appear to be reasonably well upheld when conditions are maximized as in fast writing.

Speed and Fine Motor Control Characteristics of Good and Poor Writers Among the Bright and Average Groups

An examination was made of speed and the force variation ratios of the four most legible and four least legible writers, selected on the basis of their total legibility scores for the five conditions of writing. These scores are reported in Table 56. All of the best writers were girls; all but one of the poor writers were boys. Seven of the eight subjects were from the average ability group, the one exception being subject number three from the bright group.

The legibility scores of the two groups are sharply differentiated. With one exception, the second normal sample of subject 13, the legibility of all

TABLE 56

LEGIBILITY, SPEED, AND FINE MOTOR CONTROL (FVR) SCORES FOR SELECTED GOOD AND POOR WRITERS AMONG BRIGHT AND AVERAGE CHILDREN

	Subj. No.	N	F	B	N^2	R	Total	Mean
GOOD WRITERS								
High Legibility								
HF	3	3.51	3.22	3.53	3.46	3.61	17.33	3.47
AF	11	3.01	2.90	3.27	3.25	3.09	15.54	3.11
AF	12	3.12	2.62	3.06	3.36	3.11	15.27	3.05
AF	15	3.06	2.93	3.15	2.70	2.88	14.72	2.94
Total		12.72	11.67	13.01	12.77	12.69	62.86	
Mean (T/N)		3.18	2.92	3.25	3.19	3.17		
Speed								
HF	3	30.5	20.3	37.0	31.7	31.2	150.7	30.1
AF	11	29.1	21.7	38.0	27.3	28.9	145.0	29.0
AF	12	25.8	21.3	26.9	24.3	22.7	121.0	24.2
AF	15	48.0	33.7	48.2	46.2	46.9	223.0	44.6
Total		133.4	97.0	150.1	129.5	129.7	639.7	
Mean (T/N)		33.4	24.3	37.5	32.4	32.4		
FVR								
HF	3	1.37	2.83	2.22	2.20	1.10	9.72	1.94
AF	11	1.65	1.83	1.51	1.56	1.00	7.55	1.51
AF	12	2.75	2.43	2.25	2.14	1.86	11.43	2.29
AF	15	1.05	1.24	1.01	1.03	.90	5.23	1.05
Total		6.82	8.33	6.99	6.93	4.86	33.93	
Mean (T/N)		1.71	2.08	1.75	1.73	1.22		
POOR WRITERS								
Low Legibility								
AF	13	2.18	2.15	1.92	2.72	2.42	11.39	2.28
AM	18	1.51	1.80	1.70	1.57	1.56	8.14	1.63
AM	19	2.47	2.24	2.51	1.67	2.14	11.03	2.21
AM	20	1.52	1.51	1.84	1.53	1.43	7.83	1.57
Total		7.68	7.70	7.97	7.49	7.55	38.39	
Mean (T/N)		1.92	1.93	1.99	1.87	1.89		
Speed								
AF	13	34.2	23.5	32.3	32.4	34.3	156.7	31.3
AM	18	39.5	30.0	39.3	34.8	34.4	178.0	35.6
AM	19	34.3	26.0	47.5	28.5	34.9	171.2	34.2
AM	20	36.1	28.2	41.7	29.4	36.0	171.4	34.3
Total		144.1	107.7	160.8	125.1	139.6	677.3	
Mean (T/N)		36.0	26.9	40.2	31.3	34.9		
FVR								
AF	13	1.72	2.07	1.20	1.22	1.35	7.56	1.51
AM	18	2.18	2.01	1.77	1.64	2.21	9.81	1.96
AM	19	1.97	3.68	1.67	2.93	2.09	12.34	2.47
AM	20	2.08	2.13	1.70	1.74	1.44	9.09	1.82
Total		7.95	9.89	6.34	7.53	7.09	38.80	
Mean (T/N)		1.99	2.47	1.59	1.88	1.77		

samples in all writing conditions is higher for the high legibility group than for the low legibility group. The sample of highest legibility in both groups was produced most frequently under the condition of best writing, and next most frequently as the second normal sample. Subject number 3, the consistently most legible writer, wrote her most legible sample under the relaxed condition. On the other hand, subject 18 in the poor handwriting group, actually produced his most legible sample under the fast writing condition.

As a group, the high legibility writers were on the average the faster writers under all conditions but the second normal trial. This difference between the good and poor writers is accounted for entirely by the very fast writing of three of the four best writers. Each of these three fast good writers produced, for example, fast samples written faster than any of the poor writers. By way of contrast, subject 15 in the top group consistently produced the most slowly written samples under each condition among both good and poor writers. The difference in total writing time between the good and poor writers is accounted for largely by the relatively slower writing time in the poor group in the first normal and the relaxed writing conditions.

The fine motor control of the good writers was on the average superior to that of the poor writers as the lower mean and total force variation ratios indicate. This average differential favoring the good writers held for all conditions except that of "best" writing. Noteworthy is the performance of subject 15. Under every condition of writing she exhibited the lowest FVR of all subjects in both groups and exercised this fine motor control while writing at the slowest rate of writing relative to the total group for each sample.

If the data for these dichotomized groups are examined by conditions, the following general observations may be made:

1. The high quality first normal samples were written on the average faster and with less force variability (greater fine motor control) than the poor samples.

2. Under fast writing conditions,

   a. legibility decreased on the average more so for good writers than for poor writers, though the good writers remained the distinctly better writers;

   b. all subjects wrote faster than their normal; and

   c. three out of four subjects in each group increased their force variability.

3. Under best writing conditions,

   a. subjects produced their most legible or second most legible sample in three out of four instances; and

   b. wrote more slowly and with lowered force variability, with one exception, than during their first normal.

4. The second normal writing condition produced samples which

   a. in only half of the instances in each group were more legible than their first normal;

   b. were in all cases save one written faster than the first normal sample; and

   c. were in all cases except two written with greater force variability.

5. Under relaxed writing conditions,

   a. the legibility of the samples was higher in two out of four cases in both groups than that for the first normal sample;

   b. writing speed tended to approximate that of the first normal sample, being relatively slower in the poor writing group; and

   c. force variability was on the average less among the good writers and in all instances less for this group than that during any of their previous writing samples. This was true of only one of the poor writers.

Regardless of the writing rate, the highly legible samples were generally produced with less force variability, suggesting a higher degree of fine motor control under varied conditions of writing. Fast writing conditions induced less disruption of fine motor control among good writers than among poor writers. Good writers were also able to maintain a high level of legibility under relaxed conditions while displaying their highest level of fine motor control. Such findings suggest that these good writers, all girls, have developed a highly

coordinated motor set in handwriting that is flexible and adaptable, yet relatively stable. The poor writers, three boys and one girl in this sampling, do not appear to have achieved a corresponding level of stability and adaptability to varied conditions of the handwriting act.

Between Condition Differences in GSR, EMG, FVR, and Legibility in Intra-Individual Comparisons

In attempting to determine if the ordering of the four psychophysiological variables plus speed and legibility,[32] for the five handwriting conditions might be a chance ordering, the Friedman Two-Way Analysis of Variance for ranked data was used. A summary of the results of this analysis for each of the six measures is given in Table 57. The analysis was done separately for the boys and girls in

TABLE 57

SUMMARY OF FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE OF RANK ORDER FOR EACH OF THE SIX FACTORS

Psychophysiological Measures, Legibility, and Speed	Children of Average and High Intelligence				Slow Learners		
	Boys (N=10)		Girls (N=10)		Boys and Girls (N=10)		
	X^2	P	X^2	P	X		P
Legibility	9.72	<.05	11.79	<.01			
GSR	4.62	<.30	9.93	<.02	3.03	<.50	>.30
Neck EMG	3.69	<.30	11.16	<.02	7.83	<.05*	
Arm EMG	7.56	<.10	20.52	<.001	3.12	<.50	>.30
Speed	24.24	<.001	21.69	<.001			
FVR	8.76	<.05	17.01	<.001	5.02	<.19	>.158**

*9 subjects.
**4 subjects.

the combined average and high I.Q. group levels, but for combined data on boys and girls in the slow learning group. These adjustments were dictated by differences in the writing task (most of the slow learning children could neither complete the writing task nor produce cursive writing) and by the necessity of insuring an equal number of cases in each category of the analysis.

[32]Legibility was ranked from the subject's least (rank 1) to highest (rank 5) to be consistent with the ordering of other variables.

If the five per cent level of significance is accepted, it will be noted first of all that the order of ranks for the bright and average girls differed significantly from chance for each of the six measures. In the case of the bright and average boys, the order of ranks in three of the six measures--legibility, arm EMG, and FVR--similarly differed significantly from chance. Thus in nine out of twelve orderings of the variables under consideration, there appears to be a marked consistency in the responses as these children moved through five successive conditions of writing. The ordering of the factors of legibility, speed, and fine motor control (FVR) by conditions are significant for both boys and girls. On the other hand, the boys appear less stable with respect to specific and general muscular tension (neck and arm EMG) and affect (GSR). Thus the bright and average girls as a group appear to have developed a more harmonious pattern of psychophysiological mechanisms well in accord with their legibility than have the boys in this ability grouping.

In the slow learning group it was not possible to obtain meaningful legibility or speed scores, while FVR's could be calculated for only four children who produced cursive writing. Of those scores which were obtainable in the areas of affect and muscular tension, the ordering of only one--general muscular tension (neck EMG)--appeared to be stable and to differ significantly from chance. It would thus seem that as a group, these children did not favorably compare in the development of stable and integrated psychophysiological mechanisms appropriate to the production of effective handwriting.

While the analysis of variance indicated that the individual ordering of the five responses for the five writing conditions could not in the majority of instances be attributed to chance (slow learners excluded), this analysis gave no clue concerning where these differences might lie. Hence a frequency tabulation was prepared showing the number of subjects falling into each rank for each of the five writing tasks. These data are shown in Table 58, grouped according to

TABLE 58

FREQUENCY TABULATION OF CASES BY RANKS FOR EACH OF THE WRITING CONDITIONS GROUPED ACCORDING TO FOUR PSYCHOPHYSIOLOGICAL MEASURES AND LEGIBILITY FOR CHILDREN OF AVERAGE AND HIGH INTELLIGENCE

Measurement	Task	Boys Rank Order 1	2	3	4	5	Girls Rank Order 1	2	3	4	5
GSR	Normal[1]	2	4	1	1	1	0	1	5	3	1
	Fast	2	3	3	0	1	7	2	0	1	0
	Best	3	0	1	4	1	3	1	2	1	3
	Normal[2]	1	2	3	2	2	0	1	3	3	3
	Relaxed	2	2	3	2	1	0	4	3	1	1
EMG-arm	Normal[1]	2	4	3	1	0	0	4	4	2	0
	Fast	6	2	1	0	1	9	1	0	0	0
	Best	1	4	2	2	1	0	2	2	6	0
	Normal[2]	1	0	3	5	1	1	2	3	2	2
	Relaxed	0	2	1	2	5	0	1	1	1	7
EMG-neck, shoulder	Normal[1]	1	5	2	2	0	2	1	3	2	2
	Fast	5	2	1	1	1	5	3	2	0	0
	Best	1	3	2	2	2	1	0	4	5	0
	Normal[2]	3	2	2	3	0	2	5	0	2	1
	Relaxed	0	1	4	2	3	0	1	1	2	6
FVR	Normal[1]	1	5	0	2	2	2	4	2	1	1
	Fast	5	2	3	0	0	7	2	0	1	0
	Best	2	0	3	4	1	0	2	2	5	1
	Normal[2]	1	3	2	2	2	1	2	4	2	1
	Relaxed	1	0	2	2	5	0	0	2	2	6
LEGIBILITY* (least to best)	Normal	1	1	4	4	-	0	3	4	3	-
	Fast	4	5	0	1	-	6	4	0	0	-
	Best	0	2	5	3	-	2	1	3	4	-
	Normal[2]	5	2	1	2	-	2	2	3	3	-

*Normal[2] sample not ranked in this analysis.

psychophysiological measures and legibility and have been tabulated separately for bright and average boys and girls. By inspection it can be seen that for all measures there is a strong tendency for the cases to cluster in ranks 1 and 2 during fast writing with relatively few appearing in ranks 4 and 5. For example, in the case of the boys, 22 of the 49 possible ranks for fast writing were rank 1 (evidence of the individual's maximum response) and 36 of 49 were in ranks 1 or 2. Conversely, only 3 of the possible 49 appeared in rank 5. With the girls, the trend appears even more pronounced with 34 of the possible 50 ranks clustered

in rank 1 and 46 in ranks 1 and 2, with none appearing in rank 5. Almost the reverse may be noted in relaxed writing with the cases clustering largely in ranks 4 and 5. Few occur in ranks 1 and 2 for either sex, except in the case of boys' relaxed writing. In fact, in relaxed writing there were no girls who showed a maximum response in any of the measures (rank 1) and in only 3 of 40 possible cases for the boys. It is evident from this that for the individual subject, maximum psychophysiological responses tended to occur during fast writing with minimum responses during relaxed writing, although the evidence for legibility in relaxed writing is not so clear-cut. Similarly, differences in frequencies of ranks for the other writing tasks do not reveal a consistent trend. As indicated in the analysis of variance, the evidence for bright and average boys is less conclusive than in the case of their counterparts among the girls, with the shoulder-neck EMG and the GSR data for boys not falling so neatly into place.

Summary and Findings of the Study

This investigation has examined certain physiological and motor aspects of behavior in relation to handwriting legibility in ten bright, ten average, and ten slow learning fifth and sixth grade children, equally divided as to sex, as the conditions of handwriting were systematically varied. For this purpose, the galvanic skin response (GSR) was used as an indicator of affective change; electromyographic recordings (EMG's) from the arm were used as indicators of specific muscle tension and similar recordings from the shoulder-neck region were used as indicators of general muscle tension; and an index of variation in point pressure recordings. The force variation ratio (FVR) was used as a measure of fine motor control.

The children were asked successively to write a standard sentence, "The quick brown fox jumps over the lazy dog." (1) in their usual way, (2) as fast as they could, (3) in their best handwriting, (4) again in their usual way, and (5) in just as relaxed a way as they could (neck and shoulders loose). It was

hypothesized that certain of these conditions might be mildly stressful, whereas others by nature of the task would minimize stress. In point of fact, none of the conditions could be construed to be a truly stress situation which imposed a threat to the children. The investigators actually dealt with an experimental situation in which differences in the conditions of the task, while present, were not intentionally maximized.

The major findings of the study are these:

1. In relative writing speed by conditions, the boys and girls of average intelligence and the bright girls performed similarly, but boys of high intelligence wrote more slowly on the average and were less homogeneous in writing rate. An analysis of variance for writing rates showed that for the combined bright and average groups, only the conditions of writing, not sex or intelligence, were significant at the one per cent level. Intercorrelations among writing times for task conditions for the combined groups ranged from .719 to .965. These substantial correlations revealed a high degree of consistency in writing speed for the individual and suggest that a set for speed may be an important factor in the handwriting act.

2. The bright and average group produced on the average their most rapid writing sample under the fast writing condition; and conversely in all but one instance (the first sample of bright girls) produced on the average their slowest writing sample under the best writing condition.

3. The degree of fine motor control, as evidenced by the force variation ratio, was least for the mean scores of the bright and average group under the condition of fast writing. The greatest degree of fine motor control occurred under relaxed writing in three of the four sub-groups, being only slightly less than the average boys' group mean for best writing. An analysis of variance of the force variation ratio data showed that fine motor control is significantly affected by writing condition at the one per cent level. To a lesser degree,

at the five per cent level, there appeared to be an interaction between sex and condition of writing, suggesting that boys and girls reacted differently in fine motor control as the conditions of writing were varied. For the combined group, all intercorrelations among task condition for force variation ratios were positive. Correlations among force variation ratios for fast, best, and the second normal writing conditions were .85 or higher and suggest a motor set that was far more consistent for the individual for these conditions than for the first and last of the five conditions of writing.

4. For legibility in the bright and average groups, an analysis of variance clearly showed that at the one per cent level of significance the girls as a group were the more legible writers. Less significant differences, at the five per cent level, indicated that legibility is a function of the condition of writing. Intercorrelations among legibility ratings by task conditions ranged from .632 to .932, showing substantial individual stability in legibility among the various conditions of writing.

5. A moderately high inverse rank order correlation of -.55 was found between legibility and time, indicating that high legibility tended to be accompanied by a slow rate of writing and vice versa. This correlation represents by no means a perfect degree of correspondence, for an examination of legibility in relation to the writing times of selected good and poor writers showed that some good writers wrote rapidly while some poor writers wrote very slowly.

6. An analysis of the legibility of individual samples showed that the most legible sample for an individual was distributed rather equally among three conditions of writing: best, normal, and relaxed. A similar type of analysis indicated that in the majority of cases as legibility increased, the force variation decreased, and vice versa. This indicated some positive degree of relationship between legibility and fine motor control as defined in this part of the study.

7. Comparisons of the performance of the four most legible writers with that of the four writers of lowest legibility among bright and average subjects revealed that the best writers as a group tended to sustain their superior legibility under all conditions while writing either faster or slower than the poor writers, and showed superior fine motor control under most conditions, most especially during relaxed writing. In short, the superior writers appeared to have developed a stable motor set in handwriting which enabled them to adapt to varying conditions in the handwriting act without serious deterioration in handwriting quality. By contrast, the poor writers were much less stable and less adaptable.

8. The within individual ordering of the magnitudes of the four psychophysiological responses for the five writing conditions showed significant coefficient of concordance (Kendall w) in 10 of the 20 subjects of average and high intelligence, whereas this occurred in only one subject in the low I.Q. group. However, of the 10 children in which the agreement was significant, seven were girls and three were boys.

9. Intra-individual comparisons of the rank ordering of legibility, affective response (GSR), specific muscular tension (EMG-arm), general muscular tension (EMG-neck), speed, and fine motor control (FVR) for the various conditions of writing were made for all subjects in the bright and average groups, while in the slow learning group only GSR's and EMG's were obtained for all subjects, plus FVR's for four subjects who produced cursive writing.

a. A Friedman Two-Way Analysis of Variance yielded these results for boys and girls separately in combined bright and average groups and for the total slow learning group:

(1) At the five per cent level of significance or better, the ordering of ranks for bright and average girls differed from a chance ordering in each of the six measures. Similarly, for bright and average boys, better than chance

ordering of ranks occurred in three of the six measures--legibility, arm EMG, and FVR. An integrated pattern of the measured psychophysiological functions in relation to handwriting legibility appeared to be more complete, and more stable as the conditions of handwriting were varied, for bright and average girls than for bright and average boys.

(2) While no one measure revealed a consistent difference in rank ordering for the different conditions of writing for each of the ability groupings, the boys and girls of the bright and average group were highly similar in consistency of performance as conditions of writing varied with respect to legibility, muscular tension in the arm (arm EMG), and fine motor control (FVR). Slow learning children were consistent in the direction of their reactions under successive conditions in only one of four measured variables; namely, in general muscular tension (neck EMG).

b. Analysis of the frequency of cases by ranks and by conditions revealed the following:

(1) For the bright and average group, maximum psychophysiological reactions accompanied by lower legibility, strongly tend to occur during the condition of fast writing. This trend is more evident for girls than for boys in this grouping.

(2) For the bright and average group, minimum psychophysiological reactions markedly tend to occur during the condition of relaxed writing, although this is accompanied by improved legibility in only about half of the cases.

(3) A consistent trend for other conditions of writing--the first and second normal and for best writing--is not evident.

(4) With respect to the rank-order patterning of the measures of GSR, arm EMG, and shoulder-neck EMG, there appear to be no observable differences in the rank patterns between bright, average, and slow learning children.

10. The relatively periodic nature of affective (GSR) activity noted in a preliminary study was confirmed--a peak immediately prior to writing, a second peak shortly before the beginning of the writing task gradually declining in amplitude, and a third sharply accelerating peak just after the cessation of writing and followed by relatively rapid deceleration.

In general, the investigation has shown for this group of sixth grade children, most of whom by this time might be thought to have developed some stability in their handwriting patterns, that their handwriting activity was definitely affected by the five conditions under which they were asked to write. Significant changes were noted in speed of writing and in fine motor control as conditions of writing in bright and average boys and girls were altered. Differences approaching significance were likewise secured for legibility. For this group, no significant differences were found for sex or intelligence in speed of writing, but the legibility of the girls' handwriting was significantly better than that of the boys. The effects of conditions of writing upon fine motor control were shown to be different for boys than for girls in the bright and average group, but for neither fine motor control nor legibility was this found to be significantly related to level of intelligence. Analysis of individual samples with respect to trends in legibility and fine motor control in the five writing conditions showed a positive relationship between these factors in the majority of cases. Selected good writers, writing either at relatively fast or relatively

slow rates, showed a more stable pattern of fine motor control while producing superior handwriting than did poor writers. As a combined group, the bright and average boys and girls possessed much greater stability in their pattern of measured psychophysiological activity and in legibility than did the slow learning children, while in the former group the girls on the average showed consistently more stability and integration of the handwriting-producing mechanism than did the boys.

CHAPTER V

THE RELATIONSHIP BETWEEN HANDWRITING PERFORMANCE AND INDEPENDENT MEASURES OF MOTOR CONTROL

Preliminary Research

Almost from its inception, the Committee on Basic Skills has been interested in the study of fine motor control in relation to the handwriting act since it would appear that such control is necessary for good handwriting quality. Certain findings on one index of fine motor control, the force variation ratio. based upon pressure variation during the handwriting process, were reported in the preceding chapter. It is the present purpose to report here data on measures of fine motor control gathered indeptently of the handwriting act.

In the 1955 study[1] of adult handwriting, two paper and pencil tests of fine motor control were devised and later discarded because sufficiently reliable scoring methods could not be devised. Instead, a task was substituted in which the subjects drew four straight, one-inch lines at a steady pace and pressure while point pressure was recorded electronically. Force variation ratios were then calculated for each of the segments. High reliability of the task was demonstrated by the substantial correlation of .88 between the sums of the FVR's of the first and third segments with those for the second and fourth. A correlation of .81 was likewise obtained between the force variation ratios for the straight line task and those of the normal handwriting samples. The magnitude

---

[1]Theodore L. Harris and G. Lawrence Rarick, Pressure Patterns in Handwriting (Madison: Committee for Research in Handwriting, Department of Education, University of Wisconsin, 1955). Later published as "Problem of Pressure in Handwriting," Journal of Experimental Education, XXVI (December, 1957), 151-78.

of the latter correlation led the investigators to conclude that similar factors were operating in the two task situations, that steadiness in the application of force was being measured, and that the force variation ratio employed was a useful index of fine motor control in handwriting.

In the field study of children and adolescents for which data were gathered in 1955-56,[2] a second fine motor control test battery was devised which required the subjects to draw a variety of straight and curved lines, eight in all, as point pressure was recorded. The intercorrelations for a sampling of the force variation ratios were consistently high, and the sum of the FVR's for two of the most highly correlated tasks, an ascending and a descending curved line, was selected as the measure of motor control. Correlations were then secured between the force variation ratios for the curved line motor control task and those for two words in the handwriting sample, "fox jumps." These correlations were significant ($P < .01$) at the fourth and sixth grade levels and slightly less significant ($P < .05$) at the tenth grade level.

It was thus evident in these two studies that there is some correspondence in fine motor control defined as pressure variation for the handwriting and the line-drawing tasks when the same methods of recording and analyzing the data of both tasks are used. A desire to explore more practicable ways of measuring fine motor control prompted two other attempts to study other dimensions of fine motor control.

The first of these attempts took place during the exploratory phases of the current project. A special apparatus was devised to secure a new measure of fine motor control. The motor control apparatus consisted of a series of 25 small disks set in an irregular pattern in a plastic slide. These disks, which were two millimeters in diameter and were placed about 3/16 of an inch apart, formed

---

[2]Theodore L. Harris and G. Lawrence Rarick, "Relationship Between Handwriting Pressure and Legibility of Handwriting in Children and Adolescents," Journal of Experimental Education, XXVIII (September, 1959), 65-84.

the pattern shown on the top of Fig. 6. A metal stylus connected to an electrical counter recorded the number of successive contacts.

In this motor control test, called a "tapping" test, each subject was given two trials after a demonstration by the experimenter and a preliminary warm-up trial which was not counted in the scoring. The directions were: "We would like you to show us how fast and accurately you can tap with this pencil." The score for the motor control test was derived according to the formula:

$$\frac{\text{Hits} - \text{Misses}}{\text{Time (Seconds)}}$$

The motor control tapping test was administered to the previously described group of bright, average, and slow learning children in the spring of 1957 when these children were in the third and fourth grades. Results of the analysis of the combined motor control scores for the two trials indicated a "t" test difference at the five per cent level of significance between the bright group and the slow learning group, the bright group of boys and girls making the superior scores. However, serious technical problems were encountered in the instrumentation such as the audible clicking of the counter and the tendency of the circuit to "spark" and record a hit instead of a near miss. Due no doubt in part to the malfunctioning of the apparatus, it was also not possible to establish a satisfactory degree of inter-trial reliability. As a result of these difficulties, the tapping test of fine motor control was discarded. A similar test pattern, however, was retained and employed in constructing a paper and pencil motor control test in which the subjects were asked to connect a series of circles under specified conditions.

The Circle Motor Control Test

The circle motor control test consisted of a series of small circles arranged in the pattern shown in Fig. 6. Subjects were asked to connect with a continuous line the circles in each of the groups marked off by parallel vertical lines. The task was repeated for each of the four rows of circles, making

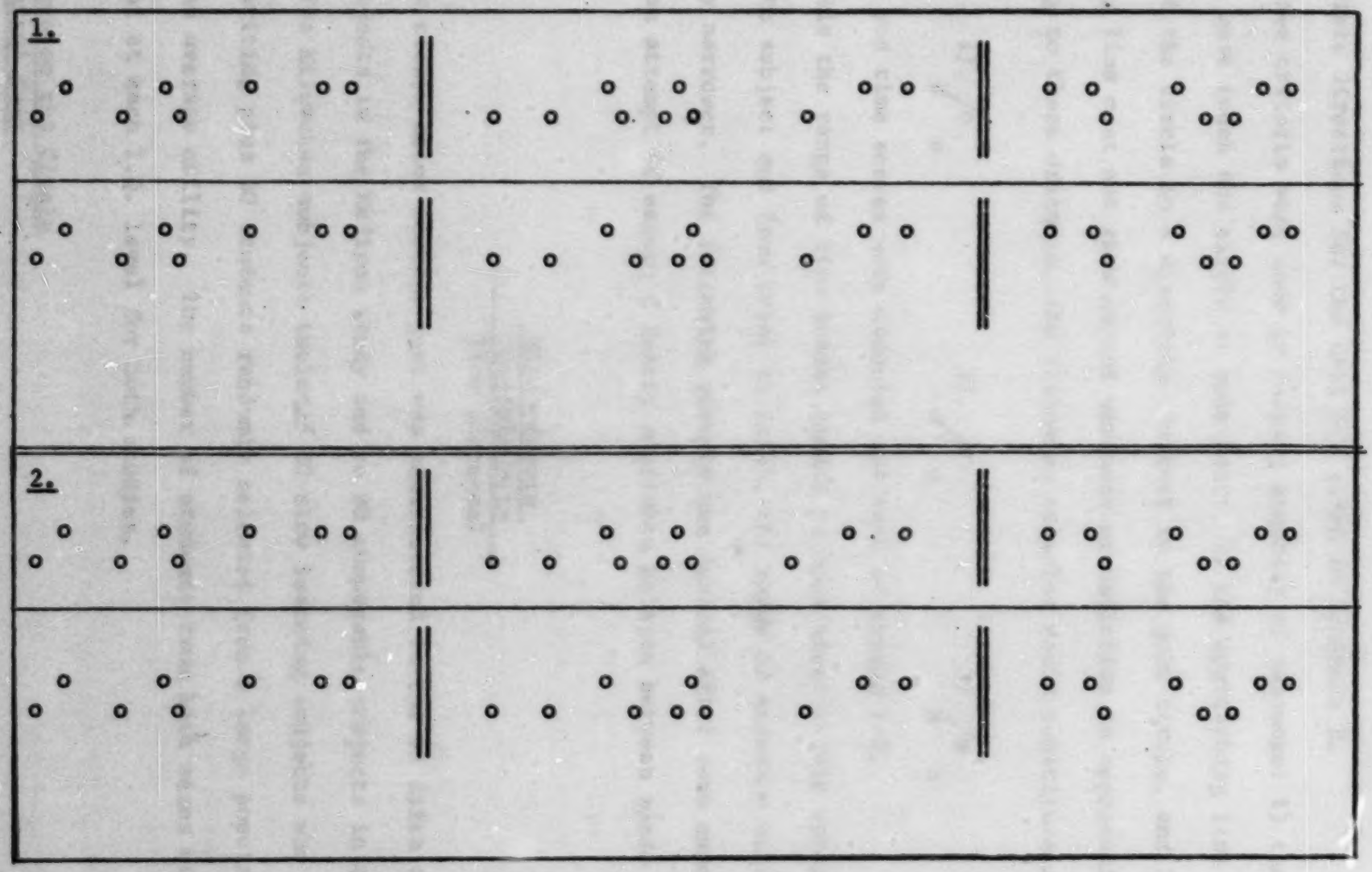


Fig. 6. Circle Motor Control Test.

four trials in all. For the first two trials subjects worked at their own speed, but for the last two trials they were instructed to connect each circle as fast as they could. Time to the nearest 1/10 of a second was recorded by stop watch. The complete directions for the test are given in Appendix E.

Three criteria were used in judging accuracy of response: 1) the approaching line must touch the circle at some point, 2) the approaching line must not overshoot the circle in a direction contrary to the next circle, and 3) the approaching line must not show marked waviness or deviation in approaching a circle. According to these criteria, the following examples would constitute errors:

1) 2) 3)

Accuracy and time scores were computed for each of trials 1-4.

While the range of time scores tended to show considerable variation from subject to subject and from trial to trial, the range of accuracy scores was considerably narrower. The following formula was devised after some experimentation in an attempt to secure a fairly equitable balance between speed and accuracy factors:

$$\frac{\frac{\text{Hits} - \text{Errors}}{\text{Possible Hits}}}{\text{Time (Seconds)}}$$

The circle motor control test was administered to the 30 fifth and sixth grade subjects in the Madison study and to 90 comparable subjects in the Milwaukee study. The Milwaukee subjects included 30 slow learning subjects who produced cursive writing plus 60 students randomly selected from a large population of bright and average ability. The number of students from both sexes was likewise equivalent at each I.Q. level for both studies.

Performance on the Circle Motor Control Test

The motor control data were treated by analysis of variance with the results of the analysis on the 90 Milwaukee subjects shown in Table 59. It will be noted

TABLE 59

RESULTS OF THE ANALYSIS OF VARIANCE ON THE CIRCLE MOTOR CONTROL TEST FOR 90 MILWAUKEE SUBJECTS

Source	df	SS	MS	F
IQ	2	154.49	72.25	15.96**
Sex	1	9.37	9.37	1.94
IQ x Sex	2	1.06	.53	
Ss/Groups	84	406.95	4.84	5.44**
Trials	3	164.35	54.78	61.55**
Trials x IQ	6	29.91	4.99	5.60**
Trials x Sex	3	1.43	.48	
Trials x Sex x IQ	6	4.35	.73	
Ss x Trials x Grps	252	223.07	.89	
	359			

**Significant at the .01 level.

that although there were no true sex differences, significant differences ($P < .01$) were found among I.Q. levels for performance upon trials 1 - 4. The large F-ratio for trials shows clearly the differential effect of the two conditions of task performance as well as possible learning effects. Reference to the progressive and consistent increase in mean scores in Table 60 from trial 1 to trial 2 and from trial 3 to trial 4 support this position. The substantial difference between the average of trials 1 - 2 and trials 3 - 4 indicate differential effects of task conditions upon motor control scores. The significant interaction of trials by I.Q. provides evidence of differential effects of trial for the three I.Q. groups. Interestingly, examination of the means and standard deviations in the test shown in Table 60 indicate that the highest mean scores were consistently obtained by the average group, next highest by the bright group, and the lowest by the slow learning group. The only exception to this pattern is in the superior performance of the bright boys in trial 4. The scatter of scores, as shown by the standard deviations, tends to be much greater on the speeded trials 3 - 4, especially for girls, than for trials 1 - 2. In the average and high I.Q. groups, the girls made higher mean scores in all trials than did the boys. In the low I.Q. group the girls were superior on three of the four trials. As pointed out earlier, however, significant sex differences were not found. Improvement in mean scores

TABLE 60

MEANS AND STANDARD DEVIATIONS ON CIRCLE MOTOR CONTROL TEST FOR 90 MILWAUKEE SUBJECTS

Group	Trial 1		2		Ave. 1-2		3		4		Ave. 3-4	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Bright Girls	3.40	.95	4.51	.98	3.96	.89	5.89	1.50	5.83	1.91	5.86	1.52
Boys	3.19	.94	4.06	1.16	3.63	.93	5.44	1.44	5.74	1.39	5.59	1.19
Total	3.30		4.29		3.78		5.67		5.79		5.73	
Average												
Girls	4.57	.95	5.22	1.00	4.89	.84	6.13	1.82	6.13	2.26	6.13	1.87
Boys	4.13	1.01	4.77	1.42	4.45	1.11	5.81	1.76	5.46	1.93	5.64	1.56
Total	4.35		4.99		4.59		5.97		5.80		5.88	
Slow Learning												
Girls	2.83	1.16	3.89	.95	3.43	.93	4.24	1.59	4.25	1.52	4.25	1.43
Boys	3.35	.82	3.61	.92	3.48	.79	3.68	.95	3.78	1.05	3.73	.88
Total	3.09		3.75		3.42		3.96		4.01		3.85	

tended to be most marked between trials 1 and 2 and between trials 2 and 3.

The results of the analysis of variance for the circle motor control test for the 30 Madison subjects are given in Table 61. Differences significant at

TABLE 61

RESULTS OF THE ANALYSIS OF VARIANCE ON THE CIRCLE MOTOR CONTROL TEST FOR 30 MADISON SUBJECTS

Source	df	SS	MS	F
IQ	2	29.1479	14.5740	4.3846*
Sex	1	10.7461	10.7461	3.2330
IQ x Sex	2	2.4656	1.2328	
Ss/Cells	24	79.7728	3.3239	2.1085**
Trials	3	2.4813	.8271	
Trials/IQ	6	7.2386	1.2064	
Trials/Sex	3	1.9688	.6563	
Trials/Sex/IQ	6	5.4590	.9098	
Trials/Sex/Cells	72	113.5043	1.5764	
	119			

*Significant at the .05 level.
**Significant at the .01 level.

the five per cent level were found between I.Q. levels but differences among trials were not significant. Examination of the means given in Table 62 showed that the high group of boys obtained the highest scores with the average and slow learning groups following in that order. Although clear-cut differences between sexes were not found, the F ratio of 3.23 approaches the five per cent level of significance. The direction of the difference favored the boys for in nine out of 12 instances the boys outperformed the girls. Contrary to the steady improvement shown between trials 1 and 2, and between trials 2 and 3 shown by the Milwaukee group, the Madison subjects as a group failed to display a consistent similar pattern of improvement on the four successive trials.

The data for the Milwaukee and Madison subjects are consistent in indicating a difference of some significance in performance on the circle motor control test by ability groups. As a group, the slow learning students in both cities performed the least well on this task. Inspection of the mean performance scores, however, indicated certain inconsistences: 1) the top performing Milwaukee group

TABLE 62

MEANS AND STANDARD DEVIATIONS ON CIRCLE MOTOR CONTROL TEST FOR 30 MADISON SUBJECTS

Group	Trial 1		Trial 2		Ave. 1-2		Trial 3		Trial 4		Ave. 3-4	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Bright Girls	4.07	.76	4.77	1.72	4.42	1.12	3.95	2.73	4.97	.87	4.46	1.33
Boys	4.44	.50	5.45	1.03	4.95	.62	4.10	1.58	4.77	1.87	4.44	1.45
Total	4.26		5.11		4.68		4.03		4.87		4.45	
Average												
Girls	3.74	1.15	4.18	.94	3.96	.88	3.12	2.12	2.97	.97	3.05	1.36
Boys	4.33	1.07	4.18	.92	4.26	.95	4.59	2.24	4.74	1.85	4.67	1.94
Total	4.03		4.18		4.11		3.86		3.86		3.86	
Slow Learning												
Girls	3.00	1.11	3.12	.49	3.06	.76	3.03	1.07	3.13	1.71	3.08	1.09
Boys	3.66	.71	3.05	.71	3.35	.57	3.91	1.83	3.98	1.17	3.95	1.42
Total	3.33		3.08		3.21		3.47		3.56		3.51	

was the average I.Q. group, whereas the bright group achieved the highest mean scores in the Madison group; 2) the Milwaukee girls tended to outperform the Milwaukee boys, while the reverse was true for the Madison subjects; and 3) significant differences between trials were obtained for the Madison group only. While perhaps greater confidence may be held for the Milwaukee data due to the large size of the sample, these inconsistencies suggested the need for a closer examination of the circle motor control test data particularly with respect to inter-trial consistency on the test.

Inter-Trial Consistency

In order to determine individual consistency by trials, intercorrelations between trials 1 and 2, 3 and 4, and the average of trials 1 - 2 and 3 - 4 were secured. These are shown for the Milwaukee subjects in Table 63. Only a fair

TABLE 63

INTER-TRIAL CONSISTENCY ON THE CIRCLE MOTOR CONTROL TEST FOR 90 MILWAUKEE SUBJECTS

Group	Trials 1-2	Trials 3-4	Average 1-2:3-4
Bright			
Girls	.722	.583	.693
Boys	.577	.418	.230
Average			
Girls	.468	.680	.664
Boys	.658	.441	.608
Slow Learning			
Girls	.638	.698	.718
Boys	.675	.535	.716

degree of consistency in performance was found between trials 1 and 2, attempted at the subject's own pace, and between the speeded trials 3 and 4. With one exception, the correlations between the average of trials 1 - 2 and trials 3 - 4 were above .60.

Intercorrelation data for the combined ability groups for the 20 Madison subjects are shown in Table 64. The inter-trial consistency on the whole tends

TABLE 64

INTER-TRIAL CONSISTENCY ON THE CIRCLE MOTOR CONTROL TEST FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	Trials		
	1-2	3-4	Average 1-2:3-4
Girls	.463	.205	.119
Boys	.499	.618	.345

to be relatively low.

It is apparent that individual performance on successive trials on the circle motor control test, especially for the Madison subjects, was highly inconsistent and unreliable. Since this test involved a ratio between speed and accuracy, the performance times on the test were separately analyzed in order to determine the specific role of speed in this task.

It will be noted in Table 65 that the inter-trial consistency for perform-

TABLE 65

INTER-TRIAL CONSISTENCY FOR TIME SCORES ON THE CIRCLE MOTOR CONTROL TEST FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	Trials		
	1-2	3-4	Average 1-2:3-4
Girls	.937	.872	.710
Boys	.907	.879	.947

ance times on the motor control test is relatively high for the bright and average Madison groups. This suggests that subjects who performed relatively fast on trials 1 - 2 likewise did so on trials 3 - 4. Thus individual consistency in

speed scores was much higher than that for scores based upon an accuracy-speed ratio on the circle motor control test reported in Table 64.

The inter-trial consistency for speed in motor control among the Madison bright and average subjects compares favorably with their inter-trial consistency in writing speed, Table 66, and in legibility, Table 67. It would appear that a

TABLE 66

INTER-TRIAL CONSISTENCY IN WRITING SPEED
FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	N^1 - F	F - B	B - N^2	N^2 - R
Girls	.923	.719	.879	.927
Boys	.787	.741	.927	.965

TABLE 67

INTER-TRIAL CONSISTENCY IN LEGIBILITY
FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	N^1 - F	F - B	B - N^2	N^2 - R
Girls	.892	.861	.722	.682
Boys	.917	.932	.771	.828

rather consistent set for speed is reflected in speed of writing and in speed scores on the circle motor control task while the subjects maintained a relatively consistent position in legibility in the five task conditions of writing.

Similar data on inter-trial consistency for the Milwaukee subjects are shown in Table 68 for writing speed and in Table 69 for legibility. As a whole these correlations do not reflect so high a degree of individual inter-trial consistency as do those for the Madison group. The lower inter-trial correlations obtained in the Milwaukee sample may in part be a function of the reduced range in writing time resulting i the homogenious grouping of subjects. It should be noted that all correlations are positive except for speed on the fast-best sample of the slow learning boys. Some evidence is thus offered that a

TABLE 68

INTER-TRIAL CONSISTENCY IN WRITING SPEED FOR 90 MILWAUKEE SUBJECTS

Group	N^1 - F	F - B	B - N^2
Bright			
Girls	.428	.235	.753
Boys	.614	.683	.907
Average			
Girls	.489	.592	.710
Boys	.714	.688	.750
Slow Learning			
Girls	.921	.765	.755
Boys	.584	-.638	.155

TABLE 69

INTER-TRIAL CONSISTENCY IN LEGIBILITY FOR 90 MILWAUKEE SUBJECTS

Group	N^1 - F	F - B	B - N^2
Bright			
Girls	.682	.580	.852
Boys	.308	.336	.745
Average			
Girls	.242	.298	.745
Boys	.703	.676	.752
Slow Learning			
Girls	.852	.719	.636
Boys	.461	.399	.796

rather consistent set for speed is operating which affects the performance of these subjects in both cities on handwriting and fine motor control tasks within the context of consistent individual handwriting legibility from trial to trial.

Relationship Between Motor Control and Legibility

Since the between-trial reliabilities of the motor control test for all groups were low, it was believed best to use the average of trials 1 and 2 and

the average of trials 3 and 4 as the performance scores for the two test conditions in examining the relationship between scores on the circle motor control test and legibility. These correlations are reported in Table 70 for each of

TABLE 70

CORRELATIONS BETWEEN AVERAGE SCORES ON TRIALS 1 - 2 AND 3 - 4 ON THE CIRCLE MOTOR CONTROL TEST AND LEGIBILITY FOR 90 MILWAUKEE SUBJECTS

Group	N^1		F		B		N^2	
	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4
Bright								
Girls	.010	.353	-.163	.261	.199	.523	.222	.410
Boys	.456	.226	.304	-.369	.489	.081	.360	.053
Average								
Girls	-.045	.024	.571	.638	-.242	-.091	-.169	-.097
Boys	.163	.165	.029	.135	-.147	.218	-.155	.136
Slow Learning								
Girls	-.258	.110	-.174	.195	.138	.421	-.010	.289
Boys	.234	-.038	.297	.310	.047	-.163	.006	-.316

the four conditions of writing for the 90 Milwaukee subjects. While 33 of the 48 correlations are positive, the correlations tend to be low and range from -.258 to .638. The highest number of positive correlations, 14, occurs among the bright group and the least number, 9, occurs among the average group which, it will be remembered, made the best scores on the circle motor control test.

Similar correlations for the 20 bright and average group of Madison subjects are shown in Table 71. Sixteen of the 20 correlations are positive but

TABLE 71

CORRELATIONS BETWEEN AVERAGE SCORES ON TRIALS 1 - 2 AND 3 - 4 ON THE CIRCLE MOTOR CONTROL TEST AND LEGIBILITY FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	N^1		F		B		N^2		R	
	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4
Girls	.379	.094	.229	.309	.331	.184	.459	-.178	.535	.427
Boys	.507	-.029	.282	.034	.327	.053	.563	-.002	.367	-.185

likewise tend to be low, ranging from -.185 to .563. All correlations between motor control and legibility of the fast and best trials are positive, however, as are those between the average of trials 1 - 2 and legibility. The low correlations and the inconsistent pattern of positive and negative correlations provide rather substantial evidence that the motor control test as scored here and handwriting performance as measured by legibility ratings include different aspects of fine motor control. If the correlations between the speed scores of the circle motor control test and legibility are examined, as shown in Table 72, a somewhat

TABLE 72

CORRELATIONS BETWEEN AVERAGE SPEED SCORES ON TRIALS 1 - 2 AND 3 - 4 ON THE CIRCLE MOTOR CONTROL TEST AND LEGIBILITY FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	N^1		F		B		N^2		R	
	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4
Girls	.172	.309	.144	-.074	.041	-.254	-.252	-.428	.084	-.123
Boys	-.675	-.670	-.348	-.426	-.577	-.473	-.591	-.573	-.364	-.358

different picture emerges for the Madison group. Fifteen of the 20 correlations are negative, and of these, all correlations for the boys are negative. Apparently, for the Madison boys, faster performance on the circle motor control test tended to be associated to a moderate degree with higher legibility scores.

Relationship Between Performance on the Circle Motor Control Test and the Force Variation Ratio

As noted earlier, the investigators had previously developed an index of fine motor control based upon variability in point pressure but which required laboratory equipment for precise movement. What is the relationship between these two measures of motor control?

Data were obtained only on the Madison subjects between these two measures. As shown in Table 73, the correlations reveal a conflicting but generally

TABLE 73

CORRELATIONS BETWEEN AVERAGE SCORES ON TRIALS 1 - 2 AND 3 - 4 ON THE CIRCLE MOTOR CONTROL TEST AND THE FORCE VARIATION RATIO FOR 20 BRIGHT AND AVERAGE MADISON SUBJECTS

Group	N^1		F		B		N^2		R	
	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4	Average 1-2	Average 3-4
Girls	.115	-.418	.589	-.062	.293	-.196	.257	-.168	-.045	-.339
Boys	.533	-.031	.160	-.249	.452	-.321	.303	-.380	.661	-.189

low relationship. The correlations range from -.418 to .661. It will be noted that all but one correlation for the average of trials 1 - 2 and the force variation ratio are positive, which suggests that higher motor control scores tended to be associated with less variability in pressure (or greater fine motor control) when subjects set their own pace on the circle motor control test. When subjects performed rapidly on this task, however, performance tended to be associated to a slight degree with greater variability in pressure (or less fine motor control).

Accuracy and Speed Factors in the Circle Motor Control Test

The preceding analysis and discussion of performance on a fine motor control test have been based upon data expressed as a ratio between accuracy and speed, and upon speed scores considered separately. Certain inconsistencies in the findings appeared to warrant an examination of the accuracy (hits) scores and speed (time) scores separately for the Madison and Milwaukee groups. The comparative mean scores for each trial are shown in Table 74 for the three ability groups.

One striking feature of this table is the relative comparability of accuracy scores, whether examined by sex or ability groups. The relative homogeneity of the accuracy scores suggests that accuracy was not the factor in the accuracy-speed ratio which effectively differentiated the performance of the subjects. The significant differentiating factor in the ratio appears to be speed. All

TABLE 74

COMPARATIVE MEANS OF HITS AND PERFORMANCE TIMES ON THE CIRCLE MOTOR CONTROL TEST FOR 20 BRIGHT AND AVERAGE MADISON AND 90 MILWAUKEE SUBJECTS

Group	I		II		III		IV	
	H	T	H	T	H	T	H	T
Bright Female								
Milwaukee	22.47	22.67	23.13	17.93	22.47	12.90	21.87	12.4
Madison	21.8	16.4	22.4	15.4	17.8	10.8	19.8	10.5
Bright Male								
Milwaukee	23.67	27.30	23.60	21.70	23.20	14.57	23.13	13.8
Madison	22.0	15.9	24.0	15.5	19.2	11.7	20.4	11.8
Average Female								
Milwaukee	24.53	19.87	23.47	15.75	22.47	12.32	21.8	11.53
Madison	22.0	19.2	22.0	17.3	17.8	11.4	17.6	11.7
Average Male								
Milwaukee	24.2	22.25	24.67	21.4	23.47	14.62	22.73	14.88
Madison	23.0	18.3	22.2	17.4	20.6	13.0	20.6	12.1
Slow Learning Female								
Milwaukee	22.6	28.97	23.73	22.55	22.07	19.0	21.93	18.1
Madison	23.6	29.0	23.4	26.9	20.2	19.5	20.6	20.5
Slow Learning Male								
Milwaukee	23.4	25.28	22.67	21.92	20.6	16.82	20.0	15.98
Madison	23.0	21.2	20.6	19.2	20.2	15.2	20.1	15.9

groups performed trial 2 faster than trial 1 whereas the ordering of trials did not result in a consistent trend in accuracy. True to expectations, the requested faster performance of trial 3 resulted in markedly reduced times in comparison with trial 2. However, for the two trials at the requested faster tempo, trial 4 was performed in less time than trial 3 in only seven of the 12 instances. On the other hand, the striking similarity of the mean accuracy scores by trials points up the fact that time was the major factor in accounting for the variance in the accuracy speed index. If the data on this table are compared with those of Tables 59 and 60, it will be seen that in most cases subjects improved their accuracy-speed motor control test scores on trials 3 - 4 while writing from one-fourth to one-third faster than on trials 1 - 2.

As a group, the slow learning children performed the slowest on the motor control test, thus confirming the differences in this direction found in both the Milwaukee and Madison studies. This analysis also indicates that as a group, the Madison subjects tended to be the faster performers in four of six comparisons, whereas the Milwaukee subjects tended to be the more accurate performers in a like number of comparisons.

Summary

Four measures of fine motor control have been described which the investigators have employed in an attempt to ascertain the relationship of motor control to various aspects of handwriting performance. Two preliminary studies utilized comparable methods of recording and analyzing the data. Each of the first two employed techniques for the precise measurement of variations in the application of point pressure, calculated as the force variation ratio (FVR). The two later studies involved tasks in which ratios between accuracy and speed were computed. One of these, a tapping test, required the subjects to hit successive small targets under unspeeded and speeded task conditions respectively. The final test, the results of which have been analyzed at some length in this chapter, was

described as a circle motor control test in which the subjects connected a series of small circles, likewise under unspeeded and speeded task conditions.

Results of the Two Preliminary Studies

In an initial study in 1955, 19 selected good and poor adult writers drew four straight parallel lines as the motor control task at a controlled pace with directions to maintain steady pressure. A correlation of .81 was obtained between the force variation ratios for this task and those of the normal handwriting samples.

In 1955-56, in a study of 144 children, 48 each at grade levels 4, 6, and 10, the motor control task of drawing two curved lines, one ascending and one descending, was used. Correlations between the force variation ratios on this task and those for two words, "fox jumps," in the normal handwriting sample were significant ($P < .01$) at grades 4 and 6 and slightly less significant ($P < .05$) at grade 10.

Results of the Tapping Test

The tapping test was administered in 1957 to the 30 children of the Madison sample when these children were in grades 3 and 4. While the slow learning children appeared to perform less well ($P < .05$) on this task than did the bright children, unsatisfactory inter-trial reliability casts some doubt upon this finding. Malfunctioning of the recording apparatus led to a discontinuance of the test in the 1959 study.

Results of the Circle Motor Control Test

This test comprised the major independent measure of fine motor control studied in the current investigation in relation to handwriting performance. Scores on this test were computed according to the formula:

$$\frac{\frac{\text{Hits} - \text{Errors}}{\text{Possible Hits}}}{\text{Time}}$$

The results of the study with respect to the performance on the circle motor control test of 90 Milwaukee children and 30 Madison children, representing bright, average, and slow learning levels of intelligence in grades 5 and 6, are now summarized.

Performance on the Circle Motor Control Test

1. An analysis of variance of performance on the circle motor control test revealed significant differences among I.Q. groups (Milwaukee, $P < .01$; Madison, $P < .05$) but no significant sex differences.

2. Mean performance on the circle motor control test were lowest in the slow learning groups except for the speeded trials 3 - 4 for the Madison boys, but differences between the bright and average groups were neither great nor consistent.

3. The speeded trials 3 - 4 tended to produce more variable behavior as indicated by larger standard deviations on these trials.

Inter-Trial Consistency for Motor Control, Writing Speed, and Legibility Data

1. Inter-trial correlations for trials 1 - 2, 3 - 4 and the averages of trials 1 - 2 and 3 - 4 on the circle motor control test ranged from relatively low to moderate magnitudes, clustering around .50.

2. Intercorrelations for the time scores alone for 20 bright and average Madison boys and girls on the circle motor control test were consistently high, being above .90 for trials 1 - 2 and above .85 for trials 3 - 4. Comparable data were not calculated for the Milwaukee group.

3. The relatively high consistency in the time scores on the circle motor control test was paralleled by correlations of similar magnitude among writing speed trials for the bright and average Madison subjects; for Milwaukee subjects, the inter-trial consistency in speed of writing was lower but was fairly substantial in most instances.

4. Similarly, inter-trial correlations for legibility in the bright and average Madison group were fairly high, the majority being above .80, but were somewhat lower for the Milwaukee subjects.

Relationship of Motor Control Performance to Legibility and Force Variation (FVR)

1. As motor control is defined and measured on the circle motor control test, its correlation with legibility was found to be low and inconsistent among subjects in both cities.

2. In relation to the FVR measure of fine motor control for the bright and average Madison subjects, the correlations were generally low, being positive in nine out of ten comparisons on unspeeded trials 1 and 2, but negative on all speeded trials 3 and 4.

The Role of Accuracy and Speed Factors in the Circle Motor Control Index

1. Accuracy scores (hits) and speed scores (time) were examined separately for all subjects on the circle motor control task, and the analysis showed that speed was the differentiating factor in the index while the accuracy scores remained relatively stable.

2. The general improvement in the circle motor control test scores found for trials 3 and 4 over those for trials 1 and 2 was largely a function of faster performance time rather than improvement in accuracy.

3. As a group, the slow learning children performed more slowly on the circle motor control test than did the bright and average children.

Discussion

In interpreting the results of this chapter in regard to motor control phenomena in relation to handwriting performance, it is well to keep in mind the components of the two types of motor control tests which were employed. In the preliminary experimentation, the force variation ratio was employed. The

calculation of this index permitted great precision in the recording of data which could be analyzed with a high degree of objectivity and consistency. In the force variation ratio, motor control in the sense of steadiness in the application of force is the paramount differentiating variable, not speed, Furthermore, since comparable records and methods of analysis were used in both the handwriting task and the fine motor control task when the latter is defined as steadiness in the application of force, one might expect a higher degree of correspondence than for measures which are relatively independent of each other.

By contrast, the use of the circle motor control test introduced several factors which apparently affected the experimental results. The measurement of speed and accuracy was less precise and involved much greater subjectivity than did the calculation of the force variation ratio. Speed was shown to be a significant differentiating variable in the circle motor control index, a circumstance which suggests that the appropriate balance between speed and accuracy factors in fine motor control presents a problem which has both practical and theoretical implications.

Such difficulties inherent in the nature and use of the circle motor control index no doubt contributed to its relatively low reliability. In addition, the very narrow range of accuracy scores obtained on this task apparently was instrumental in restricting the influence of accuracy in the motor control index. In view of these circumstances, it is not surprising that low and somewhat inconsistent relationships were found between this index of motor control and the factors of legibility, speed of writing, and steadiness in the application of force as measured by the force variation ratio. However, in spite of the relatively low reliability of the circle motor control index, significant differences between I.Q. levels were found. This fact suggests that the index does indeed possess some discriminating power.

It would also appear that such an index, or more especially the time scores involved in the index, clearly demonstrates the significant role of speed as a

differentiating variable in motor control. It has also been shown that a consistent set for speed appears to be common to handwriting and to the circle motor control task. This set for speed seems to hold irrespective of the level of legibility which in turn appears to be quite consistent.

The larger question which is raised in connection with the findings of this study is that of the specificity of the motor control function in tasks involving specific fine motor control abilities. At the present, the force variation ratio appears to be the most effective single measure of fine motor control in the handwriting task. It may be so simply because it measures a facet of motor control inherent in and highly specific to the handwriting act itself. Other measures of fine motor control gathered independently of the handwriting act would appear to require motor control functions highly analagous to those involved in handwriting, a broader range of accuracy scores than those obtained in this research, and greater precision in measurement. In addition, the appropriate balance between the speed and accuracy components in fine motor control as applied to the act of handwriting awaits further investigation.

This research has identified important components in motor control and has shown that the subject's motor control set, involving accuracy and speed, is highly significant in determining handwriting performance. It appears to represent a promising area for further study in which the significant variables may be brought under more precise control.

CHAPTER VI

CONCLUSIONS AND DISCUSSION

The examination of differences in the normative and aspirational perception behavior of bright, normal, and dull groups of right handed children in the intermediate grades in respect to the assigned tasks of handwriting and related motor and physiological conditions provides the basis for the following findings:

I. The measurement of legibility:

A. A research scale of handwriting legibility was developed covering the full range of legibility in the cursive handwriting of right handed fourth, fifth, and sixth grade school children, composed of 617 separate samples of handwriting, and arranged on a cardinal scale of legibility. This scale has greater discriminative power than any handwriting scale now known to be available for research use.

B. The items in the scale are measured on an objective basis for size and slant characteristics, thus enabling the selection of scale items of any given legibility in respect to known size and slant factors.

C. The basis has been laid for determining the number of judges necessary for the assignment of a known degree of legibility to a single copy of handwriting at a determined level of confidence. This degree of difference-level of confidence model for the required number of judges table would be an extremely valuable tool for research workers in this field.

II. The perception of handwriting behavior:[1]

---

[1]Conclusions in this and the following sections pertain to the findings on two populations--those in an intensive laboratory study (Madison) and those in the larger field investigation (Milwaukee)--unless otherwise noted.

A. The three intelligence groups differed in their performance in ordering handwriting samples as to quality; the performance of the bright and average groups was not significantly different in respect to each other; the bright and average groups differed significantly from the performance of the dull group. The dull group did not improve on successive orderings as the other groups did.

B. All intelligence groups placed their two orderings of samples "like mine" and aspired to samples at a level more consistent than chance. The bright group was most stable for placement of samples "like mine;" the bright and average groups were most stable for the placement of aspired to samples; the dull group was the least stable in both kinds of perceptions as measured by the consistency of ordering.

C. All intelligence groups placed their aspired to sample equal to or higher in quality than the sample "like mine."

D. All groups failed to place their best and poorest writing samples on a scale in two orderings at a level better than chance.

E. All groups did place their best and poorest samples relatively; i.e., best as having higher quality than poorest. The dull group was least consistent in its performance on this task.

III. The relationship between perception and legibility as measured by a legibility scale:

A. Legibility in handwriting is differentiated by sex, by gross levels of intelligence, and by writing conditions. Girls are consistently better writers than the boys, bright and average groups are better writers than the dull, and the best samples are judged better than the fast.

B. The bright and average groups are consistently more accurate than the dull group in ordering their own handwriting samples according to legibility. The average group is more accurate than either the bright or the dull.

C. Girls are more accurate than the boys in the bright and average groups in ordering their own handwriting samples according to legibility. In the dull group there is no sex difference.

D. There is no statistical difference in the normative perceptions of all three intelligence groups as measured by the legibility consistency of their N^1 and N^2 (usual) samples. There is some suggestion that these perceptions as measured by this behavior were more consistent in the average and bright groups than they were in the dull group. Similarly, boys tend to be more consistent than girls.

E. The bright and average groups, and girls, wrote over a wider range of quality than did the dull group, and boys.

F. The bright and average groups of children perceived their handwriting samples in a similar fashion and had a similar level of legibility; the low group, however, performed differently from the bright and average groups in both these respects.

G. The performance of three groups on the perception tasks were not different, however, when a generalized perception score was used.

IV. The relationships between legibility, intelligence, conditions of writing, fine motor control, writing speed, affective response, specific muscular tension, and general muscular tension in the laboratory study:

A. In relative writing speed by writing conditions, the boys and girls of average intelligence and the bright girls performed similarly, but boys of high intelligence wrote more slowly on the average and were less homogeneous in writing rate. For the combined groups, only the conditions of writing, not sex or intelligence, were significant at the one per cent level. For the individual, the high degree of consistency in writing speed suggests that a set for speed may be an important factor in the handwriting act.

B. Fine motor control as measured by the force variation ratio is significantly affected by writing condition at the one per cent level of confidence. To a lesser degree, at the five per cent level of confidence,

the interaction between sex and condition of writing suggests that boys and girls react differently in fine motor control as the conditions of writing are varied. For the combined group, the correlations between the fast, best, and the second normal writing (.85 or higher) suggest a motor set for these conditions for the individual which is far more consistent for these conditions than for the first and last of the conditions of writing.

C. The rank order correlation of -.55 found between legibility and rate of writing indicates that high legibility tends to be accompanied by a slow rate of handwriting and vice versa. There is far from a perfect degree of inverse correspondence, however, between speed of writing and legibility; some good writers wrote rapidly and some poor writers wrote very slowly.

D. The most legible sample for an individual was distributed rather equally among three conditions of writing: best, normal, and relaxed. The poorest sample was produced usually in the fast condition. Generally, as legibility increased, the force variation ratio decreased, and vice versa. This suggests a relationship between legibility and fine motor control.

E. Best writers as a group tended to sustain their superior legibility under all conditions while writing either faster or slower than poor writers, and showed superior fine motor control especially during relaxed writing. Superior writers appeared to have developed a stable motor set in handwriting which enabled them to adapt to varying conditions in the handwriting act without serious deterioration in handwriting quality. Poor writers were much less stable and less adaptable.

F. The within individual ordering of the magnitudes of the four psychophysiological responses for the five writing conditions showed a significant

coefficient of concordance in 10 of the 20 subjects of average and high intelligence; this was true for only one subject in the low I.Q. group. Of the 10 children in which the agreement was significant, however, seven were girls and three were boys.

G. Intra-individual comparisons of the rank ordering of legibility, affective response (GSR), specific muscular tension (EMG-arm), general muscular tension (EMG-neck), writing speed and fine motor control (FVR) were made for all subjects in the bright and average groups while only GSR's and EMG's were obtained for all subjects in the slow group plus FVR's for the four subjects who produced cursive samples in the Madison sample. These comparisons showed:

1. An integrated pattern of the measured psychophysiological functions in relation to handwriting legibility appeared to be more complete and more stable as the conditions of handwriting were varied for bright and average girls than for bright and average boys.
2. No one measure revealed a consistent difference in rank ordering for the different conditions of writing for each of the ability groupings. Boys and girls of the bright and average groups were highly similar, however, in consistency of performance as conditions of writing varied with respect to legibility, muscular tension in the arm, and fine motor control. Slow learning children were consistent in the direction of their reactions under successive conditions in only one of four measured variables: namely, in general muscular tension (neck EMG).
3. For the bright and average group, maximum psychophysiological reactions accompanied by lower legibility occur frequently during the condition of fast writing. This trend is more evident for girls than for boys in this grouping.

4. For the bright and average group, minimum psychophysiological reactions markedly tend to occur during the condition of relaxed writing, although this is accompanied by improved legibility in only about half of the cases.

5. There is no observable difference in rank patterns between bright, average, and slow learning children for the measures of GSR, arm EMG and neck EMG.

H. The relatively periodic nature of affective (GSR) activity noted in a preliminary study was confirmed--a peak immediately prior to writing while directions were being given, a second peak shortly before the beginning of the writing task gradually declining in amplitude, and a third sharply accelerating peak just after the cessation of writing and followed by relatively rapid deceleration.

V. The independent measures of motor control:

A. For fine motor control defined as steadiness in the application of point pressure,

1. The force variation ratio, computed from point pressure data as the subjects drew four parallel lines at a controlled pace, correlated .81 with the force variation ratio of handwriting.

2. A similar motor control test employing the force variation ratio in drawing two curved lines (ascending and descending) gave significant correlations with FVR's for two words, "Fox jumps," with subjects in grades 4 and 6.

3. These results provide evidence that steadiness in the application of force would appear to be characteristic of the subject in tasks which are similar, but not identical to handwriting.

B. For fine motor control defined as efficient adjustment of speed and accuracy components on the circle motor control test,

1. Significant I.Q. differences were found in fine motor control, although inspection of the means disclosed less difference between the average and high I.Q. groups than between either of these and the slow learners.
2. No significant sex differences were secured.
3. The circle motor control test showed low correlations with legibility and with the force variation ratio in handwriting.
4. The time scores by trials on the circle motor control test showed high intercorrelations, which corresponded favorably to substantial intercorrelations among trials in writing speed.

Discussion

The findings of this study may be discussed in relation to four major questions:

I. Do the data show differential results according to the intellectual level (bright, average, and dull) of the groups of children?

II. Do the children's normative and aspirational perceptions of their handwriting behavior differ?

III. Do the data reveal an integrated pattern among the motor and physiological correlates of handwriting?

IV. Do the results differ according to sex?

These questions will be examined in light of the findings of this study and of those of other similar recent studies. Following this, some attempt will be made to draw a number of implications for subsequent research and for educational practice in skill learning.

What is the Effect of the Three Differentiating Conditions of Intelligence on the Data of the Study?

It can be said with considerable confidence that the conditional variable of intelligence was associated with differences in the behavior patterns of the

three groups of children. This difference was most pronounced when the performances of the bright and average groups were compared with the performance of the dull group. These differences were least clear when the performances of the bright and average groups were compared with each other.

Bright and average groups differed significantly from the dull group:

In ordering the quality of scaled handwriting samples.

In the consistency of the two orderings of samples "like mine" and aspired to.

In writing at a higher level of legibility.

In ordering their own writing samples as to legibility.

In perception and legibility patterns.

In range of quality of writing. The bright and average groups wrote over a wider range of quality than did the dull group.

In magnitudes of the four psychophysiological measures (affective response [GSR], specific muscular tension [EMG-arm], general muscular tension [EMG-neck], and fine motor control [FVK]) for within the individual ordering. (Laboratory study).

In consistency of performance, as conditions of writing varied, with respect to legibility, muscular tension in the arm, and fine motor control. (Slow learning children were most consistent in only one of the measured variables: namely, in <u>general muscular tension</u>.) (Laboratory study.)

In performance on the circle motor control test. The dull group had the slowest performance. (Laboratory study.)

Bright and average groups consistently differed from the dull group but <u>not</u> at a high level of confidence:

In placing their best and poorest writing samples relatively.

In the legibility consistency of their two normal samples (N^1, N^2).

Bright and average groups differed in respect to each other but <u>not</u> at a high level of confidence:

In placement of "like mine" samples. The bright group was more consistent.

In accuracy of placement of own handwriting samples as to legibility. The average group was more accurate.

In the relative legibility difference of their aspired to sample and selected "like mine" sample. The average group had the least difference of quality, the bright and dull groups the greatest. All groups, however, placed the aspired to sample equal to or higher than the "like mine" sample.

For the bright and average groups (laboratory study):

Maximum psychophysiological reactions accompanied by lower legibility tended to occur during fast writing.

Minimum psychophysiological reactions tended to occur during relaxed writing with improved legibility occurring in only half the cases.

No difference between the three intelligence groups existed:

In relative placement of aspired to sample in respect to the sample "like mine." All groups placed the aspired to sample equal to or higher than the sample "like mine."

In the consistency of their two orderings of their best and poorest writing samples.

In relative placement of their best and poorest samples. All groups placed their best sample equal to or higher than their poorest sample.

In the legibility consistency of their normal samples (N^1, N^2).

In the total perception score.

In rank patterns for measures of GSR, arm EMG, and neck EMG. (Laboratory study.)

In the periodic nature of affective GSR during the act of writing. (Laboratory study.)

The examination of these somewhat conflicting findings suggest a number of generalizations about the relative effect of differences in intelligence on the performances of the three groups of children in the experimental tasks of this study.

1. The effect of intelligence on the performance data is not consistent over the range of intelligence considered in this study.

No reliable differences were found between the three intelligence groups both in relative placement of aspired to and "like mine" samples and in the relative placement of best and poorest writing samples. This was also true for

consistency measures for normal samples and in total perception scores. Children at all three intelligence levels were making reliable distinctions between the two kinds of perceptions being studied in this project. The significant differences in performance between these groups came when a degree difference of quality or consistency in the full range of these tasks was considered.

The great differences in the performances of the three groups occurred between the bright-normal groups and the dull group, not between the performances of the bright and average groups. Why this is true is not suggested by the data; but in addition to the difference in the factors which were used to define intelligence, the factors of experience and educational deprivation and related psychological concomitants both in and out of school must be considered in any adequate explanation of these findings.

It seems clear that in respect to all levels of intelligence, there is far less development than one would expect at this educational level of what appears to be critical perceptions in the area of skill learning. No individual group in this study appeared to have developed either very clear perceptions of what it would like its handwriting to look like, or even more strange, clear percepts of how their handwriting actually appears to them. It is difficult to see how increased maturity in a skill like handwriting can take place at the middle grade and upper levels without significant improvement in the crystallization and use of the child's perception of what his handwriting is and what it should be.

2. An important difference in performance of groups differentiated by intelligence seems to come at the level of a generalized set or overall organization or integration of the skill act. This difference in level of generalized set was evidenced by the greater capacity of more of the children in the bright and average groups to maintain a high legibility and fine motor control in writing while variation was being made in writing task and speed.

In the combined groups, only the conditions of writing were significant as a differentiating factor at the one per cent level, not the factor of sex or intelligence. In individuals, there was a high consistency in legibility of writing and speed. This suggests quite clearly the role of a generalized set related to legibility of writing, motor control, and speed.

3. There seems to be some difference in the general significance of the skill act for groups differentiated by intelligence as indicated by differences between levels of the aspired to and "like mine" samples and the differences among the general muscular tension indices for the three groups.

The performance of dull children in many of the tasks was more often characterized by chance or random behavior than was the performance of the other groups. The dull group as well as the bright group more often set greater differences between their aspired to sample and their "like mine" sample than the average group which seemed to be less anxious in their aspirations and placed their aspired to sample closer to their "like mine" sample in respect to legibility difference. The dull group was consistently different from the other groups in having a high general tension index (EMG's for neck) suggesting that the writing tasks had a more consistent threat value for them.

These findings are similar to those reported by Goldiamond:[2] namely, "Children with lower I.Q.'s gave consistently more over estimative responses in signal detection (yes-no), used more definite responses, gave more false alarms, reported more signal detection at intensities too low for accurate location, than did normal children, for whom such responses tended to diminish with age. Retarded children also overestimated in size consistency tasks."

Linton, Epstein, and Hartford in their study[3] of beginning strokes in the handwriting of 53 male freshmen of high intelligence at Brooklyn College reported

---

[2]Isral Goldiamond, "Visual Signal Detection, Perception, and Response Variables as Functions of Development and Mental Retardation," (Carbondale, Illinois: Southern Illinois University, 1961), i.v.

[3]H. B. Linton, L. Epstein, and H. Hartford, "Personality and Perceptual Correlates...," 280.

behavior for the subjects who used many beginning strokes which is similar in many ways to that reported for many dull groups--less intellectual ability, less active approach in dealing with new unstructured tasks, more likely to accept cliches, more likely to respond to emotional stimulation with avoidance, and more likely to experience anxiety and difficulty in impulse control. These studies further reinforce the findings of this study.

What is the Nature of the Differences Which Exist Between the Normative and Aspirational Perception of Children?

The general nature of the relationship which exists between normative and aspirational perceptions has been suggested by the discussion of the previous question. The following generalizations can be made with some confidence:

1. Children at the fourth, fifth, and sixth grade levels in our population, irrespective of level of I.Q., could distinguish between a writing sample like their own and the one to which they would aspire to.

It is of considerable theoretical importance that this is true. Any rationalization of the possible role and function of self direction and evaluation in skill development depends on the demonstration that the individual child can provide the rudiments for this role; i.e., can perceive where he is and where he wishes to go in his skill development.

This finding, in a sense, demonstrates the ability of children at the three intelligence levels of this study to project the images of their own handwriting to some future potential level of quality. Since this is true for all the intelligence levels of this study, the base is laid for a different kind of skill training than is customarily given to children at these levels. Recognition of this fact refutes the argument that slow learning children at this level cannot differentiate between degrees of quality and between perceptual models similar to those used in this study.

2. While all groups could order scale samples on a basis better than chance, it is clear that the generalized concept of legibility as a value is not well formed in the children of any group as measured by consistency of ordering "like mine" and aspired to samples, and best and poorest samples.

This suggests that attention paid to the development of this general legibility-value-set might be more effective in skill training on this level than coutinued attention to the development of the specific forms of letters and words. It is clear that four to six years of training in handwriting, based primarily on this latter kind of practice, is not producing this generalized kind of set. If it can be further demonstrated that the presence of this value set is of importance in high level maintenance of quality writing, a new and important direction in skill training will be determined.

3. The aspired to percept seemed to be more stable than the "like mine" perception.

This finding may account for why many children find it difficult to relate a preferred model of skill behavior to their own performance. What their own handwriting behavior actually is seems to be a relatively unstable perception and thus forms a fluctuating base for specific skill modification. Yet, presumably, the child is having far more experience with his own handwriting forms than he is with any other handwriting model. This finding suggests, therefore, that he never becomes fully aware of his own handwriting's true appearance.

Again, this possible explanation needs verification, but if further research bears out this proposition, it will support the new direction in skill training suggested by the other findings of this study.

4. What little evidence there is suggests that bright and average children are more likely than dull children to stabilize these percepts and that good writers have them to a higher degree of integration than do poor writers.

Both of these findings again tend to add further evidence to support the general thesis which seems to be emerging from these data. The qualitive difference in the skill act comes when there is some stabilization among the important perceptual and motor components of a skill act like handwriting. The components of this skill act which have highest potential for carrying it to a higher quality level seem to lie in the perceptual domain of the skill act rather than in the domain of automatic motor responses.

Is There A Patterning of the Motor and Psychophysiological Correlates in Respect to the Conditions of Intelligence and Legibility?

The evidence of patterning in respect to intelligence is not clear other than to suggest that well formed perceptual sets were not found in any of the three intelligence groups. The average and bright groups were the only ones which gave much evidence of a constant set in relation to aspired samples and accuracy of legibility placement.

There is a suggestion of a set for speed of writing which in turn may be controlled by some normative or aspirational set. Many individuals revealed a high degree of consistency in writing speed.

A fine motor control set as evidenced by consistency of the force variation ratio measure in the writing conditions of fast, best, and second normal samples in the experimental writing tasks seems clearly indicated. This was not true, however, for the first task (normal sample) and the last task (relaxed sample) in this series of tasks. It is quite likely that a consistent level of motor control is not evidenced in the initial and final phases of an act of handwriting. The middle phase of the writing task is more likely to be characterized by a consistent level of motor control. This finding has been substantiated in a number of other studies of pressure patterns in the handwriting act.

The negative correlation of -.55 between legibility and rate of writing suggests that some kind of normative or aspirational set is operating to control

speed in the writing of most children. Efficient writing seems to demand the highest rate consistent with the highest legibility. As has been pointed out, some good writers write rapidly and some poor writers write very slowly.

The most distinctive demonstration of a generalized writing set comes in the finding that good writers sustained superior legibility under all conditions of speed and writing tasks and poor writers were much less adaptable and stable in their writing performance.

Maximum psychophysiological conditions for bright and average children accompanied by lower legibility occurred during fast writing; minimum conditions occurred during relaxed writing but without pronounced improvement of legibility.

The motor control measures used in this study differed in several important respects, and it is not surprising that the performance on these tests showed a great deal of specificity. The circle motor control test places emphasis upon speed of performance with accuracy, whereas the force variation ratio is a measure of steadiness in the application of point pressure with no premium upon speed. Furthermore, the FVR is obtained either as the subject is writing or under conditions similar to handwriting, whereas the circle motor control test is further removed from the nature of the writing task. It should be pointed out, however, that the evidence indicates that speed is a variable which in itself appears to be common to the two tests and to handwriting as well. This then suggests that there is a speed set in fine manipulative skills which is characteristic of the individual.

While the broad range of speed scores for the circle motor control test yielded some insight into the operation of a speed set as task conditions were varied, the narrow range of accuracy scores did not clarify the presence or nature of a set for accuracy. The presence of the accuracy scores merely permits the observation that most children were reasonably accurate on this test. If further study verifies the existence of a speed set for fine manual skills such

as handwriting, however, this may have important implications in skill learning, for speed and precision are two essential components in motor performance. As this investigation has disclosed, fast writers are frequently good writers and poor handwriting is frequently produced at slow writing rates. Yet, since it is well known that slow movement differs from fast movement, it is conceivable that an optimum balance between the accuracy, or legibility, set, and that for speed in handwriting be sought, and that an earlier and greater emphasis upon the speed component in handwriting may be desirable in the habituation of the cursive handwriting pattern.

These findings all suggest the importance of some general organization within the learner of the factors of legibility value, speed, and nature of writing task as they relate to maintenance of a desirable quality of writing under these conditions. Poor writers seem to have such organization less well developed and thus are less stable and less adaptable in their handwriting behavior.

Particularly interesting, therefore, is the evidence presented here for the presence of generalized sets for the control of speed, level of writing pressure, and quality of legibility in the handwriting behavior of the children of this study.

Do the Findings Differ in Respect to Sex?

Studies which contrast normal and dull groups of children in respect to intelligence customarily do not find sex differences. In the present study, however, sex differences seem to be an important factor in many areas. Perhaps the most important difference is that girls write better generally than do boys and that all the other differences found may stem from this.

The specific findings in respect to the sex factor are:

Girls were consistently better writers than boys.

Girls were more accurate than boys in ordering their handwriting samples according to legibility; there was no difference, however, between sexes in this respect in the dull group.

Boys tended to have more stable normative perceptions (N^1 - N^2) than girls.

Girls wrote over a wider range of quality than did boys.

Boys of high intelligence wrote more slowly and were less homogeneous in writing rate than were boys and girls of average intelligence and girls of high intelligence.

Girls tended to have a more consistent pattern of fine motor control than did boys.

Girls in the average and bright groups had a more integrated pattern of motor and psychophysiological functioning in relation to legibility of writing than did bright and average boys.

No sex difference appeared in the circle motor control test. As one examines this factor of sex difference in these data and finds the evidence for better legibility, better quality perception, wider range of writing performance, more consistent and integrated patterns of motor and psychophysiological behavior for girls, the question as to whether these differences are not due to maturational difference rather than learning difference is bound to suggest itself. The fact that many of these girls are two or more years closer to many physiological maturities than boys at this period in their educational lives is likely to be an important explanation for the above finding.

A second explanation is the general social expectation that women are supposed to write better than men and that women teachers are thus more likely to place greater pressure on girls than on boys for better quality in their handwriting, thus producing greater attention, practice, and resultant stabilized patterns of perception and writing among girls.

It is recognized that the above explanation lies in the area of conjecture rather than in substantiated data. The evidence of broad firsthand experience with children and with schools suggests, however, that the role expectation for girls in the area of handwriting is different from that for boys and is likely to play some part in producing what are apparently sex differences in this area of skill behavior.

Perhaps the best documented explanation for these differences is found in what is known about developmental sex differences which will need to be examined further.

The lack of any clear sex differences in the low intelligence group is suggestive of another point of view on this question, but the data for this portion of the sample is too incomplete to explore the reason for this lack of sex difference more adequately.

Implications

Three major related implications are drawn, therefore, from the findings of this study.

I. The presence and differentiation of normative and aspirational perceptions have been demonstrated as being present in the writing behavior of all intelligence groups, bright, average, and dull. The lack of stability in the writing performance of middle grade children suggests the importance of this kind of development in the skill training of elementary school children. The importance of this finding, if verified, for redirection of skill training of elementary school children of this age level seems clear.

II. There is need for verifying research on these two perceptual models for handwriting--research which will systematically vary the normative and aspirational perceptions as independent variables against the dependent variable of legibility and under different conditions of the handwriting task.

III. The sex differences found in favor of girls need to be examined further in respect to what this implies for the dull group and for questions of maturation versus training.

The present findings in this area of handwriting strongly suggest the importance of extending this kind of investigation of perceptual sets to the related

language areas of reading and spelling. Irrespective of whether this is done, however, the present findings, if verified by subsequent research, will provide the basis for developing what appears to be a much more adequate handwriting educational program for use in the elementary schools of America.

APPENDICES

Appendix	Page
A. Judging of Legibility	185
B. Procedures for Development of Handwriting Project III	187
C. Letter Sent to Parents of Children Selected for Participation in Handwriting Studies.	191
D. Procedure for Symbol Perception Study I Directions for Experimenter and for Operator of Polygraph.	193
E. Directions for Handwriting Study: Madison, 1959	199
F. Directions for Handwriting Study: Milwaukee, 1959	205
G. Reproduction of Typical Point Pressure, GSR and EMG Curves of Handwriting	210

APPENDIX A

JUDGING OF LEGIBILITY

Criterion

Legibility. Webster defines "legible" as that which is "capable of being read; plain; easy-to-read."

Overview

1. Group samples into as many piles as can be clearly differentiated in terms of relative legibility. A minimum of five piles should be established. It is conceivable that only one sample might constitute by itself a separate "pile."

2. Representative samples will then be chosen as tentative scale points. Distinctiveness of tentative scale points will be checked by at least 10 judges by means of independent rank ordering of the samples.

3. Selected samples will be reproduced, through a special duplicating process, as points on a scale for rating the legibility of handwriting.

4. All samples (representing the particular grade level for which the scale has been constructed) will be rated in terms of the scale.

Suggestions for Grouping

1. First shuffle the entire set and set up three quality groups: good, poor, and in-between.

2. Then separate and carefully re-group each of the three groups into as many piles as seem necessary in order to have samples of relative legibility together.

3. Base decisions as to sample placement upon the paired-comparison method--i.e., examining one sample in relation to one other sample. In employing this method, it is often helpful to vary the relative position of cards: [A] [B], then [B] [A], and [A] and [B]. For comparison purposes, it is also helpful to overlap the cards so that the lines of writing may be brought as close as possible to each other.

4. Select a sample which you believe to be representative of each legibility group, and then carefully re-check each of your groupings against this selected sample, making any grouping revisions which seem necessary.

Cautions

1. Do not operate on a single bias such as slant, size, letter alignment, letter formation, density of line, etc. These factors should <u>all</u> enter into difficult decisions--but beware of basing such decisions on <u>only</u> <u>one</u> such factor.

Appendix A

2. Do not work so long that you become overtired. Fatigue may affect your ability to make legibility discriminations.

3. Guard against making hasty snap judgments under pressure of time, monotony, or fatigue. Try to examine each comparison by careful reading of the sample sentence--familiar though it may be. When necessary, compare each sentence phrase by phrase--i.e., "The quick brown fox" of sample A with "The quick brown fox" of sample B; "jumped over" with "jumped over;" and "the lazy dog" with "the lazy dog."

4. Remember your criterion is always legibility. In this connection, it is well to keep in mind that neatness and legibility may, or may not, be synonymous. Similarly, beauty and legibility may, or may not, be synonymous. Disregard spelling mistakes, word substitutions, etc.

APPENDIX B

PROCEDURES FOR DEVELOPMENT OF PROJECT III, 1958-59

Selection of Handwriting Samples

I. Sentence.
   A. The quick brown fox jumps over the lazy dog.
      1. Written on 5x8 lined cards.
      2. All right handed samples.
      3. Source of samples:
         a. grade school children (grades 2,3,4,5,6,8).
         b. high school children (10,12).
         c. college students.
         d. adults.

II. Judges and judging.
   A. Total of eight judges used.
   B. Judging.
      1. Judging began with fourth grade level and went up to grade 12 and to college level and adult level. Judges looked for items with specific size, slant, and legibility characteristics. As items meeting those characteristics were found, they were taken as scale items and that scale position was considered filled. Other items from later samples, which may have filled the position as adequately, were ignored.
      2. Since this is a composite scale for all age levels, the scale shows differences in slant and size of letters on each quality level.

Procedure in Selecting Size

I. Size.
   A. By general impression.- Judges rated selected handwriting samples, rating by immediate impression (large, medium, or small). Impression recorded on tabulation sheet.
   B. By scale (see Scale A).- Lined celluloid scale was placed over each sample to help measure the size of the writing--by height of loop and size of base letter. Judging recorded on tabulation sheet.
   C. Primary factors in judging size:
      1. Small size.
         a. base letters occupy 1/4 or less of the space on Scale A.
         b. loop occupies 1/2 or less of the space on Scale A.
      2. Medium size.
         a. base letters occupy 1/3 of the space on Scale A.
         b. loop occupies 2/3 of the space on Scale A.
      3. Large size.
         a. base letters occupy 1/2 or more of the space on Scale A.
         b. loop occupies 2/3 or more of the space on Scale A.
   D. Secondary factors in judging size:
      1. Small size.- the entire sentence occupies less than 8 inches on the 5x8 sample card (leaves margin over an inch deep on each edge of sentence).

Appendix B

2. Medium size.- the entire sentence occupies about 8 inches on the 5x8 sample card (little or no margin).
3. Large size.- the entire sentence occupies more than 8 inches on the 5x8 sample card (sample handwriting fills more than one line).

Data of the judgings were combined on one tabulation sheet. Three out of four, and four out of four agreement from the judgings were used in selecting the next item (slant).

Procedure for Selecting Slant
(Using Resultant Samples of Size Elimination)

I. Slant.
A. By general impression.- Judges rated samples as 1, 2, or 3 (see Scale B) according to immediate impression of dominant slant. Impression recorded on tabulation sheet.
B. By scale.- Celluloid scale showing 20° intervals of slant (see Scale B) was placed over each sample to help determine slant dominance.
C. Primary factors in judging slant.
1. Look at sample to find dominant slant.
2. Place scale over sample and find which line, if any, lines up with the dominant slant of the card and gives the sample the rating of the line. If the dominant slant lies between two sets of angle lines, give the sample the rating of the space.
3. If dominant slant cannot be determined easily, go through each word and examine slant to find the dominant slant.

Data of the judgings were combined on one tabulation sheet. Samples with three out of four, and four out of four agreement from the judgings were used in selecting the next item (legibility).

Legibility Scale Rating
(Using Resultant Size and Slant Samples)

I. Legibility
A. General impression.- Judges ranked samples from most to least legible by immediate impression. Five levels of legibility were 1 (best), 3 (next best), 5, 7, and 9 (to least legible).
B. Ranking against present scale.- Size and slant samples were ranked through comparison by finding the sample on the present scale most like the sample being ranked.

Data of the judgings were combined on one tabulation sheet. Three out of four and four out of four agreement samples were selected as the best sample of size/slant/legibility for each scale position until all scale positions were filled.

The following illustrates the placement of samples on the scale. (Scale itself is too large for reproduction.)

Appendix B

COMPOSITE HANDWRITING SCALE

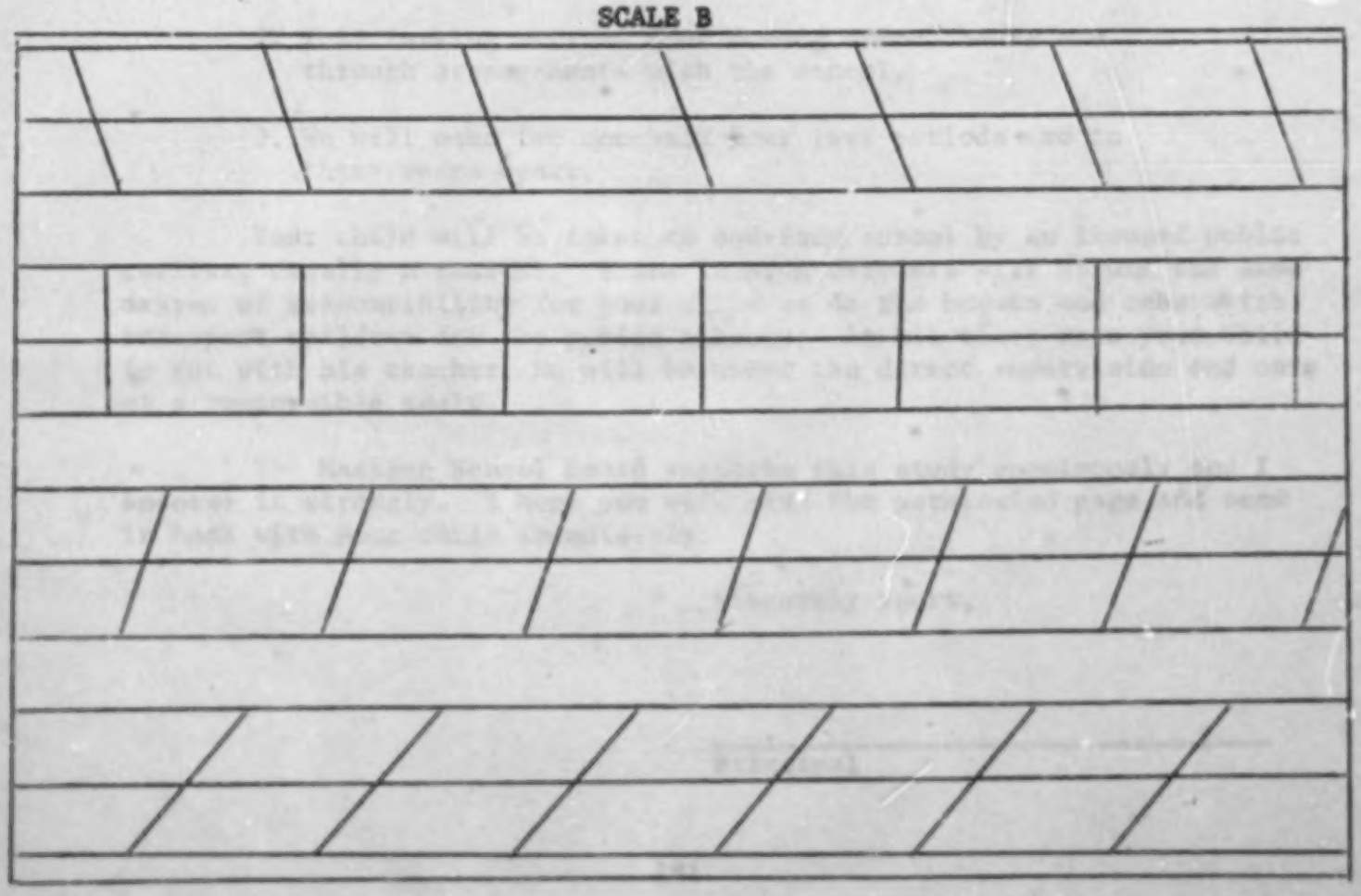
Levels of Legibility	Small	Medium	Large
1.	1 slant	1 slant	1 slant
	2 " *	2 "	2 "
	3 "	3 "	3 "
3.	1 slant	1 slant	1 slant
	2 "	2 "	2 "
	3 "	3 "	3 "
5.	1 slant	1 slant	1 slant
	2 "	2 "	2 "
	3 "	3 "	3 "
7.	1 slant	1 slant	1 slant
	2 "	2 "	2 "
	3 "	3 "	3 "
9.	1 slant	1 slant	1 slant
	2 "	2 "	2 "
	3 "	3 "	3 "

*No item found for this position.

Appendix B

SCALE A

SCALE B



APPENDIX C

LETTER SENT TO PARENTS OF CHILDREN SELECTED FOR PARTICIPATION IN HANDWRITING STUDIES

Dear Parent:

Our school with other city schools are participating in a study which we think will be helpful in getting a better understanding of how children can learn to write and do the basic skills. Professors Virgil E. Herrick and Theodore L. Harris of the University are directing the study in cooperation with us. The individual children they would like to study are selected to represent the typical groups of children attending the Madison Public Schools.

In addition to our usual records and tests given at school, the following is needed:

1. A short presentation of a writing task to the child at the University where we can keep careful records of how the child writes a short sentence. Briefly, we will want to record the pressure he exerts when he writes, his speed, his judgment of how well he is writing and his motor coordination. Children will find this interesting and enjoyable.

2. This testing will be done during school hours and through arrangements with the school.

3. We will need two one-half hour test periods two to three weeks apart.

Your child will be taken to and from school by an insured public carrier, usually a taxicab. These insured carriers will assume the same degree of responsibility for your child as do the busses and cabs which transport children for the public schools. At all times when your child is not with his teacher, he will be under the direct supervision and care of a responsible adult.

The Madison School Board supports this study unanimously and I endorse it strongly. I hope you will sign the permission page and send it back with your child immediately.

Sincerely yours,

_____

Principal

APPENDIX C

PERMISSION FORM

I hereby grant permission to _____ to take part in the study of handwriting and skill learning and to be transported to the University for two short testing periods as outlined in the attached letter.

Permission is also granted for the results of these examinations to be given to Mrs. Helen Holcombe, Assistant Director of the Child Study Department of the Madison Public Schools for use with the child and parents.

Signed_____
Parent or Guardian

APPENDIX D

PROCEDURE FOR SYMBOL PERCEPTION STUDY I
(Directions for Experimenter)

Preparatory Steps

Check needed equipment and materials:
- --Telebinocular: (1) connect (2) check light (3) place slides in order (4) check recording forms.
- --Motor coordination: (1) connect (2) set counter.
- --Randomize individual scale samples.
- --Have coded Record Forms on Hand-Symbol Perception Experiment I.
- --Code, arrange, and tape a supply of cards for handwriting samples.
- --Have Handwriting Scale available.
- --Wind stopwatch.
- --Fill and check experimental pen.

Experimenter	Directions to Subject
A. Preliminary Testing (may be done outside or inside booth).	
1. Give Telebinocular tests 3,4,5,6. Record.	1. (Follow manual instructions).
2. Give motor coordination test. Record. Situation: Table with child sitting across from the investigator. Materials: Tapping test, stop watch, and record sheet. --Be sure dots are clean and contacts are positive when stylus touches. Use steel wool if necessary. Have each child take 3 trials. One preliminary to accustom the subject to task. Then 2 trials for record. Record hits, misses, and time for each trial. (Score: Hits -number recorded on counter; Misses = 25 -hits; Time in seconds from stopwatch).	2. We would like you to show us how rapidly and accurately you can tap with this stylus (pencil). You will start here (show beginning point) and poke each dot as fast as you can in the order shown by the white line. Watch how I hold the stylus and use it to jab the dots (show child and explain the clicks and counter). You are to make only one jab at each dot.
3. Have S select desired sample from randomized scale samples. Record.	3. Choose the one you would like your handwriting to be most like.
B. Reproduction of Stimulus.	

Appendix D

Experimenter	Directions to Subject
--Take S into booth. Adjust seat and foot height. Adjust angle of platen. Tape sample card in place over platen. Have stop watch and Record Form on hand.	--Here is a little room in which we want you to show us how you write. Now let's get comfortable at this desk.
--Give S experimental pen and have him write his name.	--Try writing your name with this pen.
--Explain purpose of the exposure window.	--In a moment we will show you something in this little window to write.
--Explain and attach electrodes.	--This will not hurt you. They will help tell us how you write. Try to forget them when you write.
4a. Tell S he will see "the fox jumps" for 5 seconds only in the window and then he is to write it. (The first samples used to standardize the reproduction of the stimulus are not recorded)	--In this little window you will see the words "the fox jumps" for 5 seconds only. Look at the words carefully until they disappear. Then write "the fox jumps on the card. Write, do not print. Are you ready? Go!
b. Expose stimulus for 5 seconds. At end of 5 seconds, tell S to start writing and start stop watch. Stop memory drum.	
c. When S is finished, record time, and replace sample card. Do not tell S if reproduction is correct or not.	
d. Replace sample card and repeat steps a, b, and c until 3 successive correctly spelled samples have been produced. (If time varies greatly here, secure additional sample(s) until stable performance time is 10% from time of middle (time) sample.	--Now lets try it again. Look at the words "the fox jumps" carefully for 5 seconds until they disappear. Then write "the fox jumps" on the card. Remember--write, do not print. Are you ready? Go!
e. If S cannot produce three successive accurately spelled samples in 7 trials, accept incorrect samples and proceed to step 5.	
5. (The next three samples are to be recorded so extreme care is needed not to disturb S while writing.)	
a. Secure usual sample.	--You have shown us that you can write "the fox jumps" when you see it in

Experimenter	Directions to Subject
	the window. Now let's try it without seeing it again in the window. Write "the fox jumps" on this card just as you usually would write it.
--Get the operator's signal to start. Time sample with stopwatch.	--Ready? Go!
--When S is finished writing, complement him and place new sample on platen.	--That's fine!
b. Secure fast sample, repeat procedure in a above.	--Now we want you to write "the fox jumps" as if you were in a very big hurry.
	--Try to write faster than you just did. Ready? Go!
--When S is finished, check to see that the speed of the fast sample is 10% faster than the normal. If so, complement S. If not, replace sample on card and say:	--That's fine. or --That isn't fast enough! Let's try it again. Try hard to write faster!
(Repeat above once. Record time, etc. Then place new sample on platen and proceed to step 5c.)	
c. Secure best sample.	--Let's try once more. This will be the last one. This time try to write "the fox jumps" the very best you can. Ready? Go!
--Get operator's signal to start.	
--When S is finished writing, complement him.	--That's a very good job! We're all through with the writing now.
--Unfasten electrodes, have S move about.	--Now let's take these things off and stretch a moment.

C. Ordering of Samples. (This may be done in booth or outside booth.)

Materials needed are: fast, usual, and best samples; appropriate scale; randomized scale samples; Record Form

6. Order usual, fast, and best sample of S from best to worst.	--Here are three samples of how you just wrote "the fox jumps."
--Place 3 samples before S in random order thus: X X X	--Look at each of these. Show me which one of these three is the most poorly written. (The hardest to read.)
--Code poorest sample.	
--Code best sample.	--Show me which one is the best written. (The easiest to read.)
7. Show S the sample identified by him as best in step 6 with scale sample originally selected by S as sample "he would like his handwriting to be most like." X X	--Here is the sample which you say is your best handwriting. Now here is the sample which you said you would like your handwriting to be most like. Look at these carefully. Do you think your handwriting is: (1)

Appendix D

Experimenter	Directions to Subject
	just as good as the other? (2) Worse than the other? (3) Better than the other?
--If S says just as good as, proceed to step 8.	
--If S says worse than, ask:	--Would you like to write better than this (your sample)?
--Record responses (may need tape recorder).	--(If yes) show me what you would need to change to make yours better.
--If S says better than, ask:	--Show me how you would need to change this to make it as good as yours.
--Hand randomized samples to S and say:	--Find a sample here that you would now like your handwriting to be most like.
8. Rank poorest sample on scale, giving S his poorest sample and the appropriate scale:	--Show me which sample on this scale you think this sample of your writing is most like.
9. Rank best sample by same technique.	9. (Same as 8.)

Appendix D

PROCEDURE FOR SYMBOL PERCEPTION STUDY I
(Directions for Operator of Polygraph)

Preparatory Steps

Connect and turn on polygraph 1/2 hour before testing.
--Calibrate polygraph.
--Adust ink supply.
--Check electrode connections on instrument, electrode paste on hand.
--Check paper supply.
--Check proper paper speed.

Connect and check operation of memory drum (stimulus should not be exposed at aperture before testing starts).

Operator

A. Preliminary Testing.

Continue adjustments above.

B. Reproduction of Stimulus.

4a. When EMG and PGR electrodes are attached, recalibrate polygraph during period in which S is learning to reproduce stimulus correctly. Check baseline, especially for point pressure.

d. If the EMG and PGR recordings do not stabilize during this period, the Operator may need to signal the Experimenter for additional records before beginning step 5.

5. Signal Experimenter when you are ready to record.

a. Start polygraph to record usual sample. Code.
b. Start polygraph to record fast sample. Code. Repeat for second fast sample if signalled by Experimenter. Code, if taken.
c. Start polygraph to record best sample. Code.

--If another record is to be taken next, keep polygraph on.

--If this is the last record, turn off and disconnect polygraph and clean ink lines.

Appendix D

RECORD FORM

Name_____ Code_____

1. Vision test (Telebinocular): Pass_____Fail_____Retest results_____
2. Motor coordination test:

Prelim:	Hits_____	Misses_____	Time_____
Trial I:	Hits_____	Misses_____	Time_____
Trial II:	Hits_____	Misses_____	Time_____

3. (Using scale values)
   A. Point out the sample you find easiest to read. 1._____
   B. Is your handwriting as easy to read as this? 2._____
   C. If not, which one of these is about as easy to read as yours? 3._____
   D. If you would like to make your handwriting more readable, which one of these would you wish it to be most like? 4._____
4. Reproduction of stimulus (code time U if unsuccessful: e.g., 25U=25 seconds, if unsuccessful).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Time in Seconds															

16	17	18	19	20

5. Normal_____(time)_____(order of quality).
   Fast_____(time)_____(order of quality).
   Best_____(time)_____(order of quality).
6. Judgment of Best sample. Referent model: Comparison of sample to model.
   A. Same. Yes_____No_____
   B. Worse (your worse than). Want it better. Yes_____No_____
      Yes--change yours.
   C. Better (yours better than).
      Change in what way.
7. New desired sample (random group). 1 2 3 4 5 6
8. Placement of best sample on scale. 1 2 3 4 5 6
9. Placement of poorest sample on scale. 1 2 3 4 5 6

APPENDIX E

DIRECTIONS FOR HANDWRITING STUDY

Perception Tasks

Madison-1959

Individual arrangements for seating children will be made to meet the specific needs of the situation.

1. Please write on this card, "My name is _____," I live at _____," "I go to the _____ School." (Use separate card for this task. This writing will be used to select the type scale to use with this child.)

2. Motor Control Circle Test.
   Materials needed for testing: Sharp #4 pencil, stopwatch.

   Directions:

   Here is a card with some circles on it. Look at the top line. There are three groups of circles (point out). I want you to connect the circles in each group with a single line shifting your hand for each group of circles (point out of necessary). If you miss a circle do not stop or go back.
   a. Try to connect all the circles in each group. Work at your own speed and try to be accurate. Let's try the first line!
      Ready? Go! (Time with stopwatch and note time on record sheet.)
   b. Now do the second line in the same way.
      Ready? Go! (Time and record.)
   c. Now I am going to ask you to try the third line. This time connect each circle as fast as you can.
      Ready? Go! (Time and record.)
   d. Now do the fourth line in the same way. Connect each circle as fast as you can.
      Ready? Go! (Time and record.)

3. Motor Control Tapping Test.
   Materials: Test target, stylus, and counter. (Set at zero for each trial.) Stopwatch.

   Directions:

   Here is a plate with some dots on it. I want you to take this pencil and touch each of the dots across the plate like this. Each time you touch the dot you will hear a click. Watch how I hold the pencil and jab at the dots. If you miss a dot, do not stop or go back. Take only one poke at each dot.

   BE SURE COUNTERS ARE SET.
   Work at your own speed and try to be accurate (hit the dots).
   Ready? Go! (Time with stopwatch and note time on record sheet.)

   RE-SET COUNTERS.
   Now try it again in the same way.
   Ready? Go! (Time and record. Re-set counters.)

Appendix E

RE-SET COUNTERS.
Now I want you to try it again. This time try to touch each dot as fast as you can. Remember if you miss a dot, do not stop or go back.
Ready? Go! (Time and record.)

RE-SET COUNTERS.
Now try it once more as fast as you can.
Ready? Go! (Time and record.)

4. Perception of Handwriting Before Writing.
Materials: Scale selected most like individual's handwriting, record sheet.

Directions:

a. Now I am going to show you some handwriting samples. (Shuffle and place in random order before subject.) Show me the one which is most like yours. (Record.)
b. Reshuffle. Show me the one which you would like your handwriting to be like. (Record.)
c. Reshuffle. Now put the samples in order from the easiest to read to the hardest to read. Put the easiest to read on the top, the next easiest next, and the hardest on the bottom. (Record.)
(This is the scale ordering you will use first in the judgment series after writing.)

4A. Perception of Handwriting Before Writing (Slow Learner).
(You will note whether the child writes manuscript or cursive when he writes his name, school, and address.) If the child writes manuscript, use the series below. Cursive uses regular scale with 1,3,5 levels.

a. Manuscript scale will include three levels: high, medium, and low.
b. Only three words on the scale: "The fox jumps."
c. Check ability to reproduce stimulus (card). If he cannot write three words, write "The fox," if not two, try one, "The."

Follow same series of judgments and tests as normal and bright children in cursive, using the portion of the task he can write.

5. Show us that you can write this sentence (card with typed model placed before the child) from memory. Practice it until it can be accurately done.

FOR EXPERIMENTAL GROUP USE THE FOLLOWING:
a. Take child into booth. Convert electrodes and prepare for writing task. Prior to converting electrodes, explain procedures to child, giving him assurance that no discomforts will result. Use strain gauge pen for reading barrel pressure--recording only one segment--the thumb.
b. Check on recording equipment and calibrate strain gauges. Check photographic equipment and lighting (three seconds).
c. START CAMERA AND TEST SERIES.
(1) Rest (preparatory period) until GSR is stable.
(2) Tense fist.
(3) Rest.
(4) Tense neck.
(5) Rest (repeat 2'-5' if necessary).

Appendix E

(6) Write just as you usually write (N^1).
(7) Rest.
(8) Write very fast (F).
(9) Rest.
(10) Write your best (B).
(11) Rest.
(12) Write as you usually write (N^2).
(13) Rest.
(14) Tense neck.
(15) Rest.
(16) Write as you usually write. Try to keep neck shoulders loose (N^3).
(17) Rest (final period) until GSR is stable.

NOTE: Standardize rest periods at approximately 10 seconds--except in number 1 to allow for stabilization of GSR--continue recording 17 until GSR returns to base line.

Standardize directions for each task. "Now I want you to write (as you usually write). Ready? Begin."

Note on oscillographic record the point at which each direction is given.

FOR REGULAR WRITING GROUP USE THE FOLLOWING:
Place card before child and indicate place: (Use fountain pen--mark cards on back, not on the writing face.)

a. Write the sentence just as you usually write (N^1).
(Time speed with stopwatch and record. Indicate N^1 on card.)
GIVE NEW CARD
b. Write the sentence as fast as you can (F).
(Time and record. Mark card F)
GIVE NEW CARD
c. Write the sentence your very best (B).
(Time and record. Mark card B)
d. Write the sentence just as you usually write (N^2).
(Time and record. Mark card N^2.)

Put these samples (N^1,F,B) you have written in order from the easiest to read to the hardest to read. Record results of ordering.

Shuffle scale samples in random order and place before the subject. Show me the sample which is most like yours. Record.

Shuffle scale samples. Show me which sample you would like your handwriting to be most like. Record.

Arrange scale in the ordering developed by subject and place his best written sample beside it (ordered in 2). Show me the sample most like yours. Record.

Keep same scale arrangement and place his poorest written sample beside it. Show me the sample most like yours. Record.

Appendix E

Reshuffle scale items and ask child to rearrange items placing easiest to read at top and hardest at bottom. Record.

Be sure all cards are coded and named. Place cards and record sheet in envelope, correctly labeled.

Appendix E

RECORD SHEET (Madison)

1959 Perception Study

1. Name_____ Code Number_____

2. Motor Control Circle Test.

Trial	Nature of Trial	Time	Errors	Score
a	Own speed			
b	Own speed			
c	As fast as he can			
d	As fast as he can			

3. Motor Control Tapping Test.

Trial	Nature of Trial	Time	Errors	Score
1	Own speed			
2	Own speed			
3	As fast as he can			
4	As fast as he can			

4. Perception of Handwriting Before Writing.

Scale Selected for Use:

CURSIVE SCALE

Check one used	Sm.			Med.			Large		
	1	2	3	1	2	3	1	2	3

Check one used	Small	Medium	Large

		Value
1. One most like yours		
2. One which you would like		
3. Scale in order	1	
	3	
	5	
	7	
	9	

5. Learning to Write Standard Sample.

		Number of times sentence written for practice								
		1	2	3	4	5	6	7	8	9
Cursive	The quick brown fox jumps over the lazy dog.									
Manuscript	The fox jumps.									
	The fox.									
	The.									

Appendix E

Writing Tasks (not experimental).

	Time	Ranking	
		Sample	Rank
a. Write just as you usually write		N^1	
b. Write very fast		F	
c. Write your best		B	
d. Write just as you usually write		N^2	

Nature of Choice	Before	Now	Scale Dif.	Subject's Scale	Correct Scale
One most like mine				xxxxx	xxxxx
One you would like				xxxxx	xxxxx
Best sample	xxx	xxx	xxx		
Poorest sample	xxx	xxx	xxx		

Reshuffle and Recording of Scale.

Value	First Ordering	Second Ordering
1		
3		
5		
7		
9		

APPENDIX F

DIRECTIONS FOR HANDWRITING STUDY

Perception Tasks

Milwaukee-1959

Individual arrangements for seating children will be made to meet the specific needs of the situation.

1. Please write on this card, "My name is _____," "I live at _____," "I go to the _____ School." (Use separate card for this task. This writing will be used to select the type scale to use with this child.)

2. Motor Control Circle Test.
   Materials needed for testing:
   a. sharp #4 pencil.
   b. stopwatch.
   c. scale.
   d. sample.
   e. handwriting scale.
   f. fountain pen.

   Directions:

   Here is a card with some circles on it. Look at the top line. There are three groups of circles (point out). I want you to connect the circles in each group with a single line shifting your hand for each group of circles (point out if necessary). If you miss a circle do not stop or go back.
   a. Try to connect all the circles in each group. Work at your own speed and try to be accurate. Let's try the first line!
      Ready? Go! (Time and record)
   b. Now do the second line in the same way.
      Ready? Go! (Time and record)
   c. Now I am going to ask you to try the third line. This time connect each circle as fast as you can.
      Ready? Go! (Time and record)
   d. Now do the 4th line in the same way. Connect each circle as fast as you can.
      Ready? Go! (Time and record)

A. Perception of Handwriting Before Writing.

Select the scale most like individual's handwriting.
Materials: Scale selected most like individual's handwriting and record sheet.

1. Directions. Now I am going to show you some handwriting samples. (Shuffle and place in random order before subject.) Show me the one which is most like yours. (Record.)

Appendix F

2. Reshuffle: Show me the one which you would like your handwriting to be like. (Record.)
3. Reshuffle: Now put the samples in order from the easiest to read to the hardest to read. Put the easiest to read on the top, the next easiest next, and the hardest on the bottom. (Record.)
   (This is the scale order you will use first in the judgment series after writing.)

A. (Special) Perception of Handwriting Before Writing (Special for Slow Learners).
(You will have noted whether the child writes manuscript or cursive when he wrote his name, school, and address. If the child writes manuscript, use the manuscript scale samples.)

a. Manuscript scale will include: three levels of legibility--high, medium, and low.
b. Manuscript scale will include only three words on scale: "The fox jumps." Cursive writers will use the regular cursive scales with five levels of legibility.

Check ability to reproduce stimulus (card). If the child cannot write three words, write "The fox," if he cannot write two words, try one, "The."

Task A (Special) 1, 2, and 3.
Follow same series of judgments and tests as normal and bright children in cursive--using the portion of the task he can write.
Show us that you can write this sentence (card with typed model, "The fox jumps," placed before child) from memory. Practice it until it can be accurately done.

Circle the number of times the child wrote the sentence for practice.

If a manuscript writer cannot write three words, have him write "The fox," if he cannot write two words, try one, "The" or "fox."

Task B. Place card before child and indicate place: (Use fountain pen--mark cards on back, not on writing face.)

a. Write the sentence just as you usually write (N^1).
   (Time speed with stopwatch and record. Indicate N^1 on card.)
   GIVE NEW CARD
b. Write sentence as fast as you can (F).
   (Time and record. Mark card F.)
   GIVE NEW CARD
c. Write sentence your very best (B).
   (Time and record. Mark card B.)
   GIVE NEW CARD
d. Write sentence just as you usually write (N^2).
   (Time and record. Mark card N^2.)

Task C. Subject judges own samples after writing.
4. Put these samples (N,F,B) you have written in order from easiest to read to the hardest to read. (Record results of ordering.)

Appendix F

Task D. S judges scale samples after writing.

5. Shuffle scale samples in random order and place before subject. Show me which sample is most like yours. (Record.)
6. Shuffle scale samples: Show me which sample you would like your handwriting to be most like. (Aspired-to sample.) (Record.)

Task E. S judges his own samples against his first ordering of scale samples.

7. Arrange scale in the ordering developed by subject in items A3 and place his best written sample beside the one he chose in C4. Show me the sample most like yours. (Record.)
8. Keep same scale arrangement and place his poorest written sample beside it. Show me the sample most like yours. (Record.)

Task F. E reshuffles scale items and has S

9. Rearrange items placing easiest to read at top and hardest to read at bottom. (Record.)

Task G. If first ordering was incorrect, E places scale items in correct order and S:

10. Picks scale sample like his best writing. (Record.)
11. Picks scale sample like his poorest writing. (Record.)

Be sure all cards are coded and named. Place cards and record sheet in envelope, correctly labeled.

Appendix F

RECORD SHEET (Milwaukee)

1959 Perception Study

1. Name_____ Code Number_____

2. Motor Control Circle Test.

Trial	Nature of Trial	Time	Errors	Score
a	Own speed			
b	Own speed			
c	As fast as he can			
d	As fast as he can			

A. Perception of Handwriting Before Writing.

Scale Selected for Use:

CURSIVE SCALE

Check one used	Sm.			Med.			Large		
	1	2	3	1	2	3	1	2	3

MANUSCRIPT SCALE (Check)

☐

		Value
1. One most like yours		
2. One which you would like		
3. Scale in order	1	
	3	
	5	
	7	
	9	

Learning to Write Standard Sample.

		Number of times sentence written for practice								
		1	2	3	4	5	6	7	8	9
Cursive	The quick brown fox jumps over the lazy dog.									
Manuscript	The fox jumps									
	The fox									
	The									

Appendix F

Writing Tasks.

B.

	Time
a. Write just as you usually write	
b. Write very fast	
c. Write your best	
d. Write just as you usually write	

C.

Ranking	
Sample	Rank
N[1]	
F	
B	

	Nature of Choice	Before	Now	Scale Difference	Subject's Scale
D.	5. One most like mine				xxxxx
	6. One you would like				xxxxx
E.	7. Best sample	xxx	xxx	xxxxx	
	8. Poorest sample	xxx	xxx	xxxxx	

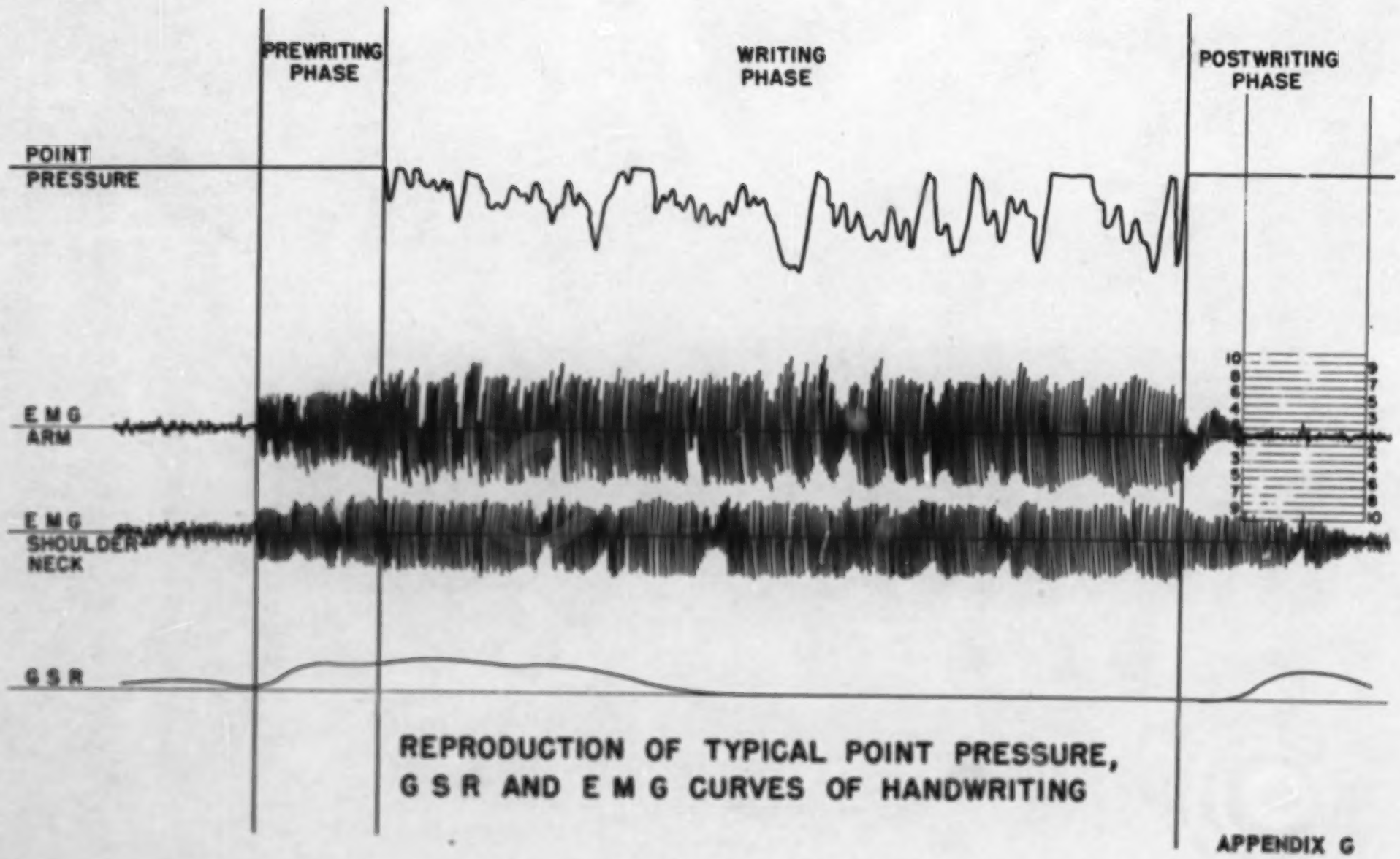
Reshuffle and Recording of Scale.

F9.

Value	First Ordering	Second Ordering
1		
3		
5		
7		
9		

G10-11.

	Using Correct Scale
Best sample	
Poorest sample	



REPRODUCTION OF TYPICAL POINT PRESSURE, G S R AND E M G CURVES OF HANDWRITING

APPENDIX G